

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 15:16:12 ; Search time 2912 Seconds
(without alignments)
3666.694 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPLDINIQEPRWD.....PAMAIPLIMDTLEKKDFLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09990415/runat_16102003_073806_9438/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415.ecgn_1_1_3745_etrunat_16102003_073806_9438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1826	6 AX061199	AX061199 Sequence
2	1345	100.0	2702	6 BD127742	BD127742 Primer fo
3	1345	100.0	2702	9 AK074707	AK074707 Homo sapi
4	1345	100.0	2735	9 BC000124	BC000124 Homo sapi
5	1345	100.0	3047	6 AX188304	AX188304 Sequence
6	1318	98.0	984	6 AX364964	AX364964 Sequence
7	1286	95.6	2699	10 AF276997	AF276997 Rattus no
8	1283	95.4	2710	10 AF325262	AF325262 Mus muscu
9	1283	95.4	2870	10 BC012208	BC012208 Mus muscu
10	1089	81.0	1134	6 AX511171	AX511171 Sequence
11	1089	81.0	1818	9 AF327346	AF327346 Homo sapi
12	1089	81.0	1839	9 AK056915	AK056915 Homo sapi
13	1089	81.0	2367	6 AX302019	AX302019 Sequence
14	1076	80.0	2236	5 BC044027	BC044027 Xenopus l
15	1064	79.1	2986	10 S70011	S70011 Rattus sp.
16	1059	78.7	1204	10 BC005743	BC005743 Mus muscu
17	1059	78.7	2803	10 AF325260	AF325260 Mus muscu
18	1053	78.3	752	6 BD124792	BD124792 Primer fo
19	1053	78.3	752	6 BD126228	BD126228 Primer fo
20	1026	76.3	2766	6 BD157602	BD157602 Primer fo
21	1026	76.3	2766	9 AK022287	AK022287 Homo sapi
22	1015	75.5	2894	9 BC020517	BC020517 Homo sapi
23	961.5	71.5	1084	6 AX457932	AX457932 Sequence
24	897.5	66.7	984	6 AX511173	AX511173 Sequence
25	897.5	66.7	984	6 AX457933	AX457933 Sequence
26	852.5	63.4	181086	9 AL133215	AL133215 Human DNA
27	798	59.3	2642	6 BD159719	BD159719 Primer fo
28	798	59.3	2642	9 AK022938	AK022938 Homo sapi
29	786.5	58.5	2062	5 BC048044	BC048044 Dario rer
30	765	56.9	741	6 BD148081	BD148081 Primer fo
31	756	56.2	1394	3 BT001832	BT001832 Drosophi1
32	754	56.1	1316	3 AY089419	AY089419 Drosophi1
33	753	56.0	1537	10 AF325261	AF325261 Mus muscu
34	753	56.0	1617	10 BC019808	BC019808 Mus muscu
35	750	55.8	1772	6 AX392949	AX392949 Sequence
36	750	55.8	2499	9 AF462052	AF462052 Homo sapi
37	750	55.8	2608	9 BC022091	BC022091 Homo sapi
38	712.5	53.0	1575	6 AX740470	AX740470 Sequence
39	711.5	52.9	25586	2 AC132957	AC132957 Mus muscu
40	705	52.4	252403	2 AC137016	AC137016 Rattus no
41	705	52.4	294396	2 AC105485	AC105485 Rattus no
42	705	52.4	347368	2 AC096620	AC096620 Rattus no
43	693	51.5	806	6 BD146776	BD146776 Primer fo
44	656	48.8	26286	2 AC121209	AC121209 Rattus no
45	650.5	48.4	229937	2 AC137328	AC137328 Rattus no

RESULT 1

ALIGNMENTS

AX061199	AX061199	1826 bp	DNA	linear	PAT 22-JAN-2001
LOCUS	Sequence 46 from Patent WO0078953.				
DEFINITION	AX061199				
ACCESSION	AX061199.1	GI:12406335			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 Lal, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O., Burford, N., Baughn, M. R., Azimzai, Y., Lu, D. A., Au-Young, J. and Patterson, C.				
TITLE	Human transport proteins				
JOURNAL	Patent: WO 0078953-A 46 28-DEC-2000;				
FEATURES	Incyte Genomics, Inc. (US)				
source	location/Qualifiers				
	1..1826				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	415 a 519 c 512 g 380 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,33e-115	Length:	1826		
Score:	1345.00	Matches:	261		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-990-415A-2 (1-261) x AX061199 (1-1826)					
QY	1 MetGluSerLysMetcLylGluLeuProLeuAspIleasnIleGlnGluProArgTyrPasp	20			
DB	449 ATGAAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGCTGGAC	508			
QY	21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu	40			
DB	509 CAAAGTACTTCTCTGGGACAGAGCCCGGCACCTTTTCACTGTTACTGATCCTCGAAATCTG	568			
QY	41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly	60			
DB	569 CTGCTGTCCGGGCGACACAGCTGGAAGCTTCTCGGACATCGTGCAAGTACAGGGCGGC	628			
QY	61 ValValThrProGlyIleThrGluAspGlnLeuTyrPArgAlaLysTyrValTyrAspSer	80			
DB	629 GTGTGACCCCGAGGATCACCGAGGACAGCTGTGTGAGGGCCAAAGTATGTGTGACTCC	688			
QY	81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal	100			
DB	689 GCCTTCCATCCGGACACACAGGGAGAGGTGTCTGATTGGCCGCATGTACAGCCAGGTG	748			
QY	101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal	120			
DB	749 CCCATGAACATGACCATCACTGCTGCATGCTCACATTCTACAGGAAGACCCCAACCGTG	808			
QY	121 ValPheTyrPglntirPValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer	140			
DB	809 GTGTCTCGGACAGTGGTGAATCAGTCCCTTCATGCCATTGTTAACACTCCAACCGCAGT	868			
QY	141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly	160			
DB	869 GGTGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCATATGTGAGTGCCACCACTGGA	928			
QY	161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly	180			
DB	929 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCAACCAAGCACCTGCCCCCTGTGTGGC	988			
QY	181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg	200			

Db	989	AGATTGTGCCCCCTTTGGCAGCAGTGGCAGCTGCCCAACTGCATCAACATCCCCCTGATGAGG	1048
Qy	201	GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyIleArgLeuGlyTyr	220
Db	1049	CAGAGAGAGCTGCAGGTGGCATCCCGGTGGCTGATGAGCAGCTCAGAGGCTTGCTTAC	1108
Qy	221	SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla	240
Db	1109	TCGGTGAAGTGCAGCCAGCAGGGAATCTTCCAGGTGGTGAATTTCAGAATCTGCATGGCG	1168
Qy	241	IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu	260
Db	1169	ATTCCCTGCCATGGCCATCCACCACCTGATCATGTGACACTCTGAGAGAAGAACTTCCTG	1228
Qy	261	Lys 261	
Db	1229	AAG 1231	
RESULT 2	BD127742	2702 bp	DNA linear
LOCUS	BD127742	2702 bp	DNA linear
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		
ACCESSION	BD127742		
VERSION	BD127742.1	GI:23222687	
KEYWORDS	JP 2002017375-A/3173.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2702) Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,T., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.		
TITLE	Primer for synthesizing full-length cDNA and use thereof		
JOURNAL	Patent: JP 2002017375-A 3173 22-JAN-2002;		
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002017375-A/3173 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII, PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA, PI TETSUJI OTSUKI,HISASHI KOGA PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS (70)..(1044).		
FEATURES	source	1.2702	
BASE COUNT	608 a 757 c 699 g 638 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3-72e-115	Length:	2702
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-990-415A-2 (1-261) x BD127742 (1-2702)			
QY	1	MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp	20
Db	70	ATGGAAGCAAAATGGGTGAATTGCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC	129

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 15:16:12 ; Search time 2912 Seconds
(without alignments)
3666.694 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPLDINIOEPRWD.....PAAIPLIMDTLEKKDLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlp
-Q=cgn2.1/USPTO_spool_p/US09990415/runat_16102003_073806_9438/app.query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415 @CGN_1_1_3745 @runat_16102003_073806_9438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1826	6 AX061199	AX061199 Sequence
2	1345	100.0	2702	6 BD127742	BD127742 Primer fo
3	1345	100.0	2702	9 AK074707	AK074707 Homo sapi
4	1345	100.0	2735	9 BC000124	BC000124 Homo sapi
5	1345	100.0	3047	6 AX188304	AX188304 Sequence
6	1318	98.0	984	6 AX364964	AX364964 Sequence
7	1286	95.6	2699	10 AF276997	AF276997 Rattus no
8	1283	95.4	2710	10 AF325262	AF325262 Mus muscu
9	1283	95.4	2870	10 BC012208	BC012208 Mus muscu
10	1089	81.0	1134	6 AX511171	AX511171 Sequence
11	1089	81.0	1818	9 AF327346	AF327346 Homo sapi
12	1089	81.0	1839	9 AK056915	AK056915 Homo sapi
13	1089	81.0	2367	6 AX302019	AX302019 Sequence
14	1076	80.0	2236	5 BC044027	BC044027 Xenopus l
15	1064	79.1	2986	10 S70011	S70011 Rattus sp.
16	1059	78.7	1204	10 BC005743	BC005743 Mus muscu
17	1059	78.7	2803	10 AF325260	AF325260 Mus muscu
18	1053	78.3	752	6 BD124792	BD124792 Primer fo
19	1053	78.3	752	6 BD126228	BD126228 Primer fo
20	1026	76.3	2766	6 BD157602	BD157602 Primer fo
21	1026	76.3	2766	9 AK022287	AK022287 Homo sapi
22	1015	75.5	2894	9 BC020517	BC020517 Homo sapi
23	961.5	71.5	1084	6 AX457932	AX457932 Sequence
24	897.5	66.7	984	6 AX457933	AX457933 Sequence
25	897.5	66.7	984	6 AX511173	AX511173 Sequence
26	852.5	63.4	181086	9 AL133215	AL133215 Human DNA
27	798	59.3	2642	6 BD159719	BD159719 Primer fo
28	798	59.3	2642	6 AK022938	AK022938 Homo sapi
29	786.5	58.5	2062	5 BC048044	BC048044 Danio rer
30	765	56.9	741	6 BD148081	BD148081 Primer fo
31	756	56.2	1394	3 BT001832	BT001832 Drosophil
32	754	56.1	1316	3 AY089419	AY089419 Drosophil
33	753	56.0	1537	10 AF325261	AF325261 Mus muscu
34	753	56.0	1617	10 BC019808	BC019808 Mus muscu
35	750	55.8	1772	6 AX392949	AX392949 Sequence
36	750	55.8	2499	9 AF462052	AF462052 Homo sapi
37	750	55.8	2608	9 BC022091	BC022091 Homo sapi
38	712.5	53.0	1575	6 AX740470	AX740470 Sequence
39	711.5	52.9	255686	2 AC132957	AC132957 Mus muscu
40	705	52.4	252403	2 AC137016	AC137016 Rattus no
41	705	52.4	294396	2 AC105485	AC105485 Rattus no
42	705	52.4	347368	2 AC096620	AC096620 Rattus no
43	693	51.5	806	6 BD146776	BD146776 Primer fo
44	656	48.8	262886	2 AC121209	AC121209 Rattus no
45	650.5	48.4	229937	2 AC137328	AC137328 Rattus no

ALIGNMENTS

RESULT 1

AX061199	LOCUS	AX061199	1826 bp	DNA	linear	PAT 22-JAN-2001
	DEFINITION	Sequence 46 from Patent WO0078953.				
	ACCESSION	AX061199				
	VERSION	AX061199.1	GI:12406335			
	KEYWORDS					
	SOURCE	Homo sapiens (human)				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	REFERENCE	1				
	AUTHORS	lal, P., Yang, J., Yue, H., Hillman, J. L., Tang, Y. T., Bandman, O., Burford, N., Baughn, M. R., Azimzal, Y., Lu, D. A., Au-Young, J. and paterson, C.				
	TITLE	Human transport proteins				
	JOURNAL	Patent: WO 0078953-A 46 28-DEC-2000; Incyte Genomics, Inc. (US)				
	FEATURES	Location/Qualifiers				
	source	1. 1826				
		/organism="Homo sapiens"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:9606"				
	BASE COUNT	415 a 519 c 512 g 380 t				
	ORIGIN					
	Alignment Scores:					
	Pred. No.:	2, 33e-115	Length:	1826		
	Score:	1345.00	Matches:	261		
	Percent Similarity:	100.00%	Conservative:	0		
	Best Local Similarity:	100.00%	Mismatches:	0		
	Query Match:	100.00%	Indels:	0		
	DB:	6	Gaps:	0		
US-09-990-415A-2 (1-261) x AX061199 (1-1826)						
QY	1	MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp	20			
Db	449	ATGGAAGCAAAATGGGTGAATTGCCTTAGACATCAACATCCAGAACCTCGCTGGGAC	508			
QY	21	GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu	40			
Db	509	CAAGTACTTCTCTGGGAGAGACCCGGCACCTTTTCACTGTACTGATCCTCGAAATCTG	568			
QY	41	LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly	60			
Db	569	CTGCTGTCCGGGGCACACGCTGGAAGCTTCTCGAATCTGTGCAGAACTACAGGCGCGC	628			
QY	61	ValValThrProGlyIleThrGluAspGlnLeuTyrPArgAlaLysTyrValTyrAspSer	80			
Db	629	GTGTGACCCCGAGGATCACCGAGACCAAGCTGTGGAGGCCCAAGTATGTGTGACTCC	688			
QY	81	AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal	100			
Db	689	GCCTTCCATCCGACACACAGGGAGAAGGTGTCTGATTTGGCCGATGTACAGCCAGGTG	748			
QY	101	PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal	120			
Db	749	CCCATGAACATGACCATCACTGGCTGCATGCTACATTTCTACAGGAAGACCCCAACCGTG	808			
QY	121	ValPheTyrGlnTyrPValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer	140			
Db	809	GTGTTCCTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTTAACTACTCCAACCGCAGT	868			
QY	141	GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly	160			
Db	869	GGTGACACTCCCATCACTGTGAGGCAGCTGGGGACAGCCCTATGTGAGTGCCCACTGGA	928			
QY	161	AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly	180			
Db	929	GCTGTGGCCACGGCCCTGGGACTCAATCCCTCAACCAAGCACCTGCCCCCTGTGGTGGC	988			
QY	181	ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg	200			

Db	989	AGATTGTGCCCCCTTTGGCAGCAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGATGAGG	1048
Qy	201	GINARGGLULEGNINVALGIYILEPROVALAIAASPGIUALAGIYGINARGLEUGLYTYR	220
Db	1049	CAGAGAGAGCTGCAGAGTGGGCATCCCGGTGGCTGATGAGGCAGTCCAGAGGCTTGGCTAC	1108
Qy	221	SERVALTHRALAALALYSGINGIYILEPHEGGINVALIISERARGILECYMETALA	240
Db	1109	TCGGTGACTGCAGCCAGCAGGGAATCTTCCAGGTGGTGATTTCAAGAATCTGCATGGCG	1168
Qy	241	ILEPROALAMETALALIEPROPROLEUILMETASPHRLEUGLULYSLSASPHLEU	260
Db	1169	ATTCCCTGCCATGGCCATCCCAACCACCTGATCATGGACACTCTGGAGAAGAACTTCCTG	1228
Qy	261	LYS 261	
Db	1229	AAG 1231	
RESULT 2	BD127742	2702 bp	DNA linear PAT 18-SEP-2002
LOCUS	BD127742		
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		
ACCESSION	BD127742		
VERSION	BD127742.1	GI:23222687	
KEYWORDS	JP 2002017375-A/3173.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2702) Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,T., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.		
TITLE	Primer for synthesizing full-length cDNA and use thereof		
JOURNAL	Patent: JP 2002017375-A 3173 22-JAN-2002;		
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002017375-A/3173 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII, PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI SHINICHI KOJIMA, PI TETSUOI OTSUKI,HISASHI KOGA PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH key		
FEATURES	FT CDS	location/Qualifiers (70)..(1044).	
source	1. 2702	location/Qualifiers	
BASE COUNT	608 a	757 c	699 g 638 t
ORIGIN			
Alignment Scores:			
Pred. NO.:	3-72e-115	Length:	2702
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-990-415A-2. (1-261) x BD127742 (1-2702)			
Qy	1	MetGluserLySmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTrpasp	20
Db	70	ATGGAAGCAAAATGGGTGAATTGGCTTTAGACATCAACATCCAGGAACCTGCTGGGAC	129


```

QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db      130 CAAAGTACTTTCCTGGGCAGAGCCCGGCACCTTTTCACTGTACTGATGATCCTCGAAATCTG 189
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db      190 CTGCTGTCCGGGGCAGACTGGAAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCCGGC 249
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db      250 GTGGTGACCCAGAGGATCACCAGGAGCAGCTGTGGAGGGCCCAAGTATGTATGACTCC 309
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db      310 GCCTTCCATCCGACACAGGGAGAGAGGTGCTCCTGATGGCCGATGTACAGCCCGG 369
QY      101 PrometasmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db      370 CCCATGAACATGACCATCCTGCTGCTGATGCTCACAATCTACAGAGAACCCCAACCGTG 429
QY      121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db      430 GTGTCTCGCAGAGTGGGTGAATCAGTCTTCAATGCCATGTGTTACTACTCCAACCGCAGT 489
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db      490 GGTGACACTCCCATCCTGATGAGGCAGCAGCTGGGGACAGCCTATGTGAGTGCCACCATGGA 549
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db      550 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACTTGCCCCCTTGGTCGGC 609
QY      181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db      610 AGATTGTGCCCCCTTGGCAGAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGG 669
QY      201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db      670 CAGAGAGAGCTGCAGGTGGCATCCCATGTGCTGATGAGGCAGGTGAGAGGCTTGGCTAC 729
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db      730 TCGGTGACTGCAGCAAGCAGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCG 789
QY      241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db      790 ATTCTGCCATGGCCATCCCACTGATCATGACACTCTGTGAGAAGAAAGACTTCTCTG 849
QY      261 Lys 261
Db      850 AAG 852

```

RESULT 3
AK074707
LOCUS Homo sapiens cDNA FLJ90226 fls, clone NT2RM1000858, weakly similar to Tricarboxylate carrier.
DEFINITION
ACCESSION AK074707
VERSION AK074707.1 GI:22760323
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.
TITLE NEDO human cDNA sequencing project

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2702)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hrl.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
FEATURES
source
1..2702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM1000858"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_1lb="NT2RM1"
/notes="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells"
70..1047
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11151.1"
/db_xref="GI:22760324"
/translation="MESKMGELPLDINIOEPRDOSTFLGRARHFTVTDPRNLLSGAOLEASRNIVQNRAGAVTPGITEDOLMRKAYVDSAFHPDTEKVVILGRMSAOVPMNMTITGCMLEFYRKPTVFWQWVNOVSNFNAIVNSNRSGDPTITVROLGTAYVSATGAVATAGLKLSTKHLPLVGRFVFPFAVAALANCLNIPLMRORELOVGIPIVADEAGORLGXSVTAAGKIFOVVISRICMAIPAMAIPLIMDTLEKDELKRPMWLGAPLQVGLVGFLVFATPLCCALFPQKSSIHISNLEPELRAQIHONPSVEVYVYINKGL"
BASE COUNT 608 a 757 c 699 g 638 t
ORIGIN
Alignment Scores:
Pred. No.: 3.72e-115 Length: 2702
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-990-415a-2 (1-261) x AK074707 (1-2702)
QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
Db      70 ATGGAAGCAAAATGGGTGAATGCTTTAGACATCAACATCCAGAACTCGCTGGGAC 129
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db      130 CAAAGTACTTTCCTGGGCAGAGCCCGGCACCTTTTCACTGTACTGATCCTCGAAATCTG 189
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db      190 CTGCTGTCCGGGGCAGACTGGAAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCCGGC 249
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db      250 GTGGTGACCCAGAGGATCACCAGGAGCAGCTGTGGAGGGCCCAAGTATGTATGACTCC 309
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db      310 GCCTTCCATCCGACACAGGGAGAGAGGTGCTCCTGATGGCCGATGTACAGCCCGG 369
QY      101 PrometasmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db      370 CCCATGAACATGACCATCCTGCTGCTGATGCTCACAATCTTACAGAGAACCCCAACCGTG 429

```

QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 430 GTGTTTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTTACTACTCCAAACCGCAGT 489
QY 141 GLYAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 490 GGTGACACTCCCATCAGTGTGAGGAGCTGGGACAGCCTATGTAGTGCCACCACTGGA 549
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 550 GCTGTGGCCACGGCCCTGGAGCTCAATCCCTCACCAAGCAGCTGCCCCCTTGTCGGC 609
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnLysIleAsnIleProLeuMetArg 200
Db 610 AGATTGTGGCCCTTTCAGACAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGG 669
QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnValGlnArgLeuGlyTyr 220
Db 670 CAGAGAGAGCTGCAGGTGGGCATCCAGTGGCTGATGAGGAGGTCAGAGGCTTGCTAC 729
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 730 TCGGTGACTGCAGCCAGCAGGAGAACTTCCAGGTGTGATTCAGAATCTGCATGGCG 789
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
Db 790 ATTCTGCCATGGCCATCCCACTGATCATGACACTCTGAGAGAAGAAGACTTCTCTG 849
QY 261 Lys 261
DI 850 AAG 852

RESULT 4
BC000124
LOCUS 2735 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, similar to tricarboxylate carrier-like protein, clone
ACCESSION MGC:3122 IMAGE:3352015, mRNA, complete cds.
VERSION BC000124
KEYWORDS BC000124.1 GI:12652746
SOURCE MGC.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2735)
Strausberg, R.
Direct Submission
Submitted (03-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

REMARK
AUTHORS
TITLE
JOURNAL
COMMENT
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 6 Row: K Column: 13
This clone was selected for full length sequencing because it

FEATURES
source passed the following selection criteria: Hexamer frequency ORF
analysis, Genomescan gene prediction.
location/Qualifiers
1. .2735

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:3122 IMAGE:3352015"
/tissue_type="Eye, retinoblastoma"
/clone_1ib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
75. .1040
/codon_start=1
/product="similar to tricarboxylate carrier-like protein"
/protein_id="AAH00124.1"
/db_xref="GI:12652747"

CDS
BASE COUNT 645 a 757 c 695 g 638 t
ORIGIN
Alignment Scores:
Pred. No.: 3.78e-115 length: 2735
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-990-415A-2 (1-261) x BC000124 (1-2735)
QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
Db 63 ATGGAAGCAAAATGGGTGAATGCCCTTTAGACATCAACATCCAGAACTCGTGGGAC 122
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 123 CAAGTACTTTCCTGGGACAGAGCCGGCAGCTTTTACAGTGTACTGATCTCGAATCTG 182

QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 183 CTGCTGTCCGGGGGACAGCTGGAAGCTTCTCGAAGATCTGTCAGAACTACAGGGCCGCG 242
QY 61 ValValThrProGlyIleThrGlnAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db 243 GTGGTGACCCAGGAGTACCCAGAGACAGCTGTGGAGGGCCAAAGTATGTATGACTCC 302
QY 81 AlaPheHisProAspThrGlyGlnLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 303 GCCTTCATCCGGACACAGGGGAGAAAGTGTCTGTGATGGCCGATGTACAGCCAGGTG 362
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 363 CCCATGAACATGACCATCTGCTGCTGCTACATCTTACAGGAAGACCCCAACCGTG 422

QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 423 GTGTTTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTTAACTACTCCAAACCGCAGT 482
QY 141 GLYAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 483 GGTGACACTCCCATCAGTGTGAGGAGCTGGGAGACCTATGTAGTGCCACCACTGGA 542

QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 543 GCTGTGGCCACGGCCCTGGAGCTCAATCCCTCACCAAGCAGCTGCCCCCTTGTCGGC 602
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnLysIleAsnIleProLeuMetArg 200

Db 49 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 108

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40

Db 109 CAAAGTACTTCTGGGCGAGAGC-CGGCACCTTTTTCACCTGTTACTGATGATCCGAAATCTG 167

QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60

Db 168 CTGCTGTCCGGGGGACAGCTGGAGCTTCGGACATCGTGACAGACTACAGGGCCGGC 227

QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80

Db 228 GTGGTGACCCCGAGGATCACCGAGACCAGCTGTGAGGGCCCAAGTATGTGATGACTCC 287

QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100

Db 288 GCCTTCATCCGGACACAGGGGAGAGGTTGCTCTGATTGGCCGATGTCAAGCCAGGTG 347

QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120

Db 348 CCCATGAACATACCATCACTGCTGCATGCTCACATTTCTACAGGAAGCCCAACCGTG 407

QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140

Db 408 GTGTTCTGGCAGTGGGTGAATCAGTCTCTCAATGCCATTGTTACTACTCCAAACCGCAGT 467

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160

Db 468 GGTGACACTCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTGAGTGCACCACTGGA 527

QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180

Db 528 GCTGTGGCCACGGCGCTGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGTCGGC 587

QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200

Db 588 AGATTGTGCCCTTTGGACAGAGTGGCAGCTGCCAAGTGCATCAACATCCCTGATGAGG 647

QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220

Db 648 CAGAGAGAGTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAGGTGAGGCTTGCTGCTAC 707

QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240

Db 708 TCGGTGACTGCAGCCCAAGCAGGAATTTCCAGGTGATGATTCAAGAATCTGCATGGCG 767

QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260

Db 768 ATTCCTGCCATGGCCATCCCAACCACTGATGACACTCTGGAGAAAGACTTCTCTG 827

QY 261 Lys 261

Db 828 AAG 830

RESULT 7

AF276997 2699 bp mRNA linear ROD 22-JUN-2000

LOCUS Rattus norvegicus tricarboxylate carrier-like protein mRNA,

DEFINITION complete cds.

AF276997

ACCESSION AF276997.1 GI:8650525

VERSION

KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 2699)

AUTHORS Mashima,H. and Kojima,I.

TITLE Direct Submission

JOURNAL Submitted (12-JUN-2000) Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15, Showa-machi, Maebashi, Gunma 371-8512, Japan

FEATURES

source location/Qualifiers

1..2699

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague Dawley"

/db_xref="taxon:10116"

233..1198

/note="CTP-like protein"

/codon_start=1

/product="tricarboxylate carrier-like protein"

/protein_id="AAF78249.1"

/db_xref="GI:8650526"

/translation="MGDLPLNINIQEPRNDSTFLGRARHFFTVTPDRNLLSGKOLEASRNIVQNRAGVVTPELTEDQLWRKAVYDSAFHPDTGEKVVLIGMSAOVPMMTITGCMLTFRKPTPTVFWQWVNSFNATVNSNRSDAPITVOQLGPAVVSATGAVATALGLSLTKHLPLVGRFVFPFAVAANANCINIPLMRQRELQVGIPVTDEAGORLGHVTAAKGIFQVQVVSRIQMAIPAMAIPVIMNTLEKDKFLKRPWLGAPLQVLGVFCLVFATPLCALFPQRSSIHVTRLEPELRAQIQAKPSIDVYYNKGIL"

BASE COUNT 611 a 747 c 684 g 657 t

ORIGIN

polyA_signal 2661..2666

Alignment Scores:

Pred. No.: 1.12e-109 Length: 2699

Score: 1286.00 Matches: 248

Percent Similarity: 98.46% Conservative: 8

Best Local Similarity: 95.38% Mismatches: 4

Query Match: 95.61% Indels: 0

Gaps: 10

US-09-990-415a-2 (1-261) x AF276997 (1-2699)

QY 2 GlnSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGln 21

Db 224 GAAAGCAAGATGGCGGCACTTGCCCTTAATATTCACATCCAGGAACCTCGATGGACCAA 283

QY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 41

Db 284 AGCACATTCCTAAGGACAGAGCCCGGCACCTTTCACAGTTACTGATGATCCCGAAATCTCCTG 343

QY 42 LeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyVal 61

Db 344 CTGTCGGGAACAGTTGGAGCTTCTCGGAACATCGTGACGAATTCAGAGGCTGTGTG 403

QY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAla 81

Db 404 GTAACCTCGTGGGCTCACTGAGGACCAAGCTATGAGAGCCAAATACGTGATGACTCAGCA 463

QY 82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101

Db 464 TTCCACCCCGGACACAGGGGAGAGGTGGTCTTGAATTGGTCCGATGTCCGCCAGGTGCC 523

QY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121

Db 524 ATGAACATGACCATTTACTGGCTGCATGCTCACCTTCTACAGGAAGACTCCCAACTGTGTG 583

QY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141

Db 584 TTCTGGCAGTGGGTGAATCAGTCTCTCAATGCTATTGTGAATTACTCTAATCGCAGTGA 643

QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161

Db 644 GACGCTCTATCACTGTGACAGCAGCTGGGAACGGCCTATGTGAGTGACCACTGGGGCT 703

QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181

Db 704 GTGGTACTGCTCTGGGACTCAAGTCTCTCACCAAGCACCTACCCCACTAGTGGGTGA 763

QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201

Db 764 TTGCTACCTTTGGACAGCTGTGGCTGGCCCAACTGCATCAACATCCCTGATGAGGCAG 823

QY 202 ArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSer 221

|||||
Db 824 AGGAGCTGCAGGTGGCATCCCACTGACTGATGAAGCTGGTCAGAGACTTGCCACTCG 883
QY 222 ValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetaIaIle 241
Db 884 GTGACTGCTGCCAAACAGGGCATCTTCCAGGTGGTATCGAGAATCGGCATGCGCATC 943
QY 242 ProAlaMetaIaIleProLeuIleMetaSphrLeuGluLysLysAspPheLeuLys 261
Db 944 CCCGCATGGCCATTCCTCCCGGTGATCATGAACACTCTGGAGAAGAAAGACTTCCTCAAG 1003
RESULT 8
AF325262
LOCUS AF325262 2710 bp mRNA linear ROD 30-APR-2001
DEFINITION Mus musculus sideroflexin 3 (Sfxn3) mRNA, complete cds.
ACCESSION AF325262
VERSION AF325262.1 GI:13785615
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2710)
Fleming, M.D., Campagna, D.R., Haslett, J.N., Trenor, C.C. III and
Andrews, N.C.
TITLE A mutation in a mitochondrial transmembrane protein is responsible
for the pleiotropic hematological and skeletal phenotype of
flexed-tail (f/f) mice
JOURNAL Genes Dev. 15 (6), 652-657 (2001)
MEDLINE 21172735
PUBMED 11274051
REFERENCE 2 (bases 1 to 2710)
AUTHORS Fleming, M.D., Campagna, D.R. and Andrews, N.C.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Pathology, Children's Hospital, 300
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source location/Qualifiers
1..2710
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2710
/gene="Sfxn3"
83..1048
/gene="Sfxn3"
/note="SFXN3; similar to the flexed-tail (f) protein
sideroflexin 1 (SFXN1)".
/codon_start=1
/product="sideroflexin 3"
/protein_id="AAK39430.1"
/db_xref="GI:13785616"
/translation="MGDLPLNINIQEPRWDSTFLGRARHFFVTDPNLLSGEQL
ASRNIVQNYRAGVATPGLTEDQLWRAKYVDSAFHPDTGEKVVLIGRMSAQPMMTI
TGCMLEFYRKPTVFWQWVNSFNALVNSNRSGADPITVQQLGTAYSATTGAVAT
ALGLSLTKHLPLLVGRFVPAVAANNCINIPIMRORELOYGIPYTDAGORLGHV
TAAKQGIQVVISRIGMAIPMAIPVIMNTLEKKDFLKRPRWICAPLQVGLVGFCLV
FATPLCCALFPQRSSIHVTRLPELRQIQANPSIDVYVYNNKGL"
BASE COUNT 591 a 755 c 647 g 711 t 6 others
ORIGIN
Alignment Scores:
Pred. No.: 2.14e-109 Length: 2710
Score: 1283.00 Matches: 248
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 95.38% Mismatches: 5
Query Match: 95.39% Indels: 0
Gaps: 0
US-09-990-415A-2 (1-261) x AF325262 (1-2710)
QY 2 GluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGln 21
Db 74 GAAAGCAAGATGGGTGACCTGCGCTTAATATCAACATCCAGAACCTCGGTGGAGACCAA 133

QY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 41
Db 134 AGCACATTTCTAGGACAGAGCCCGCATTTCTTCCACAGTCACTGATCCYCGAAATCTGCTG 193
QY 42 LeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyVal 61
Db 194 CTGTCCGGGGAACAGCTGGAGCTTCCCGGAACATCGTGCAGAAATTACAGGGCTGTGTG 253
QY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAla 81
Db 254 GCAACCCCGGTCTCACTGAGAGACAGCTATGGCGAGCCAAATACGTATGACTACACA 313
QY 82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101
Db 314 TTCCATCCGGACACGGGGGAGAGGCTGCTTGTATTGGCCGYATGTCAGCCAGGTGCC 373
QY 102 MetaSmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
Db 374 ATGAACATGACCATTTACTGCTGCATGCTCACCCTTCTACAGAGAAAGACTCCGACTGTG 433
QY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
Db 434 TTCTGGCAGTGGGTCAATCAGTCCCTTCATATGCTATGTGAATTACTCTAATCGCAGGCG 493
QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
Db 494 GATGCTCCCATCACTGTGCAGCAGTGGGAGACAGCCATATGTAGTGCCACCACCTGGGCT 553
QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
Db 554 GTGGCTACTGCTCTGGGACTCAAGTCTCTCACCACCAAGCACTGCCCTAGTGGTGA 613
QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetaGln 201
Db 614 TTCGTGCCCTTTGACGTGTGGCGSCTGCCAAGTCAACATCATCCCTGATGAGCGAG 673
QY 202 ArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSer 221
Db 674 AGGAGCTGCAGGTGGGCATCCAGTACTGATGAGGCTGTGAGAGGCTGGCCACTCG 733
QY 222 ValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetaIaIle 241
Db 734 GTGACTGCTGCCAAACAGGAATCTTCCAGGTGGTGTATCAAGAATCGGAATGGCATTT 793
QY 242 ProAlaMetaIaIleProLeuIleMetaSphrLeuGluLysLysAspPheLeuLys 261
Db 794 CCCGCATGGCCATTCCTCCCGGTGATCATGAACACTCTGGAGAAGAAAGACTTCCTGAAG 853
RESULT 9
BC012208 2870 bp mRNA linear ROD 16-APR-2003
LOCUS BC012208
DEFINITION Mus musculus sideroflexin 3, mRNA (cDNA clone MGC:18958
IMAGE:3985155), complete cds.
ACCESSION BC012208
VERSION BC012208.1 GI:15126553
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2870)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Urdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

TITLE Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Boufard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 JOURNAL MEDLINE PUBMED 12477932 2 (bases 1 to 2870) Strausberg,R. Direct Submission Submitted (06-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov COMMENT Contact: MGC help desk Email: cgaphs-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAK Plate: 23 Row: 1 Column: 8 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein. FEATURES Location/Qualifiers 1..2870 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="MGC:18958 IMAGE:3985155" /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue." /clone_11b="NCI_CGAP_Mam5" /lab_host="DH10B" /note="Vector: PCMV-SPORT6" 1..2870 /gene="Sfxn3" /db_xref="LOCUSID:94280" /db_xref="MGI:2137679" 221..1186 /codon_start=1 /product="sideroflexin 3" /protein_id="AAH12208.1" /db_xref="GI:15126554" /db_xref="LOCUSID:94280" /translation="MGDLPLNINIQEPRWDQSTFLGRARHFEVTDPNLLSGEOL EASRNIVONRAGVATPGLTEDQIMRAKVVYDSAFHPDTGEKVVILGRMSAOVPMNMTI TGCMLTFYRKPTPTVFWQWQNSFNIAIVNYSNRSGDAPITVOOLGTAYVSATGAVAT ALGKSLTKHLPLPLVGRFVFAVAANCINIPLMRQRELQVGIPTDEAGQRLGHSV TAAKQGFQVVISRIGMAIPMAIPVIMNTLEKDFLKRPRMIGAPLQVGLVGFCLV FATPLCCALFPQRSSIHVTRLPELPAQIOAQNPSIDVVIYNKGL" BASE COUNT 622 a 804 c 696 g 748 t ORIGIN Alignment Scores: 2.29e-109 Length: 2870 Pred. No.: 2870

Score: 1283.00 Matches: 248 Percent Similarity: 98.08% Conservative: 7 Best Local Similarity: 95.38% Mismatches: 5 Query Match: 95.39% Indels: 0 DB: 10 Gaps: 0 US-09-990-415a-2 (1-261) x BC012208 (1-2870) QY 2 GluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGln 21 Db 212 GAAAGCAAGATGGGTGACTGCTCCCTTAATATATCAACATCCAGGAACCTCGGTGGACCAA 271 QY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 41 Db 272 AGCACATTTCTAGGCAGAGCCCGCATTTCTTACAGTCACTGATCCCCGAAATCTGCTG 331 QY 42 LeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyraArgAlaGlyVal 61 Db 332 CTGTCCGGGGAACAGCTGGAAGCTTCCCGGAACATCGTGCAGAAATTACAGGCGTGTGTG 391 QY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyraValTyraSpSerAla 81 Db 392 GCAACCCCGGGTCTCAGTACAGGACCACTATGGCGAGCAATAAGTATGACTGACGCA 451 QY 82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101 Db 452 TTCATCCCGGACACGGGGAGAAAGGTGCTGTGATTGGCCGTATGTCAAGCCAGGTGCC 511 QY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyraArgLysThrProThrValVal 121 Db 512 ATGAACATGACCATTTACTGCGTGCATGCTCACCTTCTACAGGAAGACTCCGACTGTGGTG 571 QY 122 PheTrpGlnTrpValaGlnIserPheAsnAlaIleValaAsnTyraSerAsnArgSergly 141 Db 572 TTCTGGCAGTGGGTCAATCAGTCTTCAATGCTATGTGAATTACTTAAATCGACGGCC 631 QY 142 AspThrProIleThrValaArgGlnLeuGlyThrAlaTyraValSerAlaThrThrGlyAla 161 Db 632 GATGCTCCCATCAGTGCAGACAGTGGGACAGCCTATGTGAGTGCCACACAGCGGCT 691 QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181 Db 692 GTGGCTACTGCTGTGGACTCAAGTCTCTCACCACAGCACTGCCCGCTAGTCGTCGA 751 QY 182 PheValProPheAlaAlaValaAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201 Db 752 TTCGTGCCCTTTCAGCTGTGGCCGCTGCCAATGCATCAACATCCCCCTGATGAGCGAG 811 QY 202 ArgGluLeuGlnValGlyIleProValaAlaAspGluAlaGlyGlnArgLeuGlyTyrSer 221 Db 812 AGGAGAGCTGACAGGTGGGCATCCAGTACTGATGAGGCTGTGACAGAGGCTTGGCCACTCG 871 QY 222 ValThrAlaAlaLeuGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIle 241 Db 872 GTGACTGCTGCCAAGCAGGAATCTTCCAGGTGTGATATCAAGAAATCGAATGGCGATT 931 QY 242 ProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261 Db 932 CCCGCCATGGCCATTCCCCCGGTGATCATGAACACTCTGGAGAAAGAAAGACTTCTGAAG 991 RESULT 10 AX511171 1134 bp DNA linear PAT 27-SEP-2002 LOCUS Sequence 1 from Patent WO0246221. AX511171 AX511171 DEFINITION AX511171 VERSION AX511171.1 GI:23392062 KEYWORDS Homo sapiens (human) SOURCE Homo sapiens ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 Herath,H.M., Parekh,R.B., Rohlf,C., Terrett,J.A. and Tyson,K.L.

TITLE Proteins
JOURNAL Patent: WO 0246221-A 1 13-JUN-2002;
Oxford GlycoSciences (UK) Limited (GB)
FEATURES Location/Qualifiers
 source 1. 1134
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
BASE COUNT 303 a 268 c 273 g 287 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 7.85e-92 Length: 1134
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: Gaps: 6

US-09-990-415A-2 (1-261) x AX511171 (1-1134)

```
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 17 GGAGAACTACCAACCAACATTACATCAAGAACCTCGATGGATCAAGCACTTCATT 76
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
Db 77 GGACGAGCCATCATTTCTCTCACTGTACTGACCCCAAGAACATCTCTTAACCAACGAA 136
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 137 CAACCTCGAGAGTGCAGAAAAATAGTACATGATTACAGGCAAGGAATGTTCTCTCTG 196
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 197 CTTACAGAAATGAATTGTGGAGAGCAAAAGTACATCTATGATTGACGTTTTCATCCTGAC 256
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 257 ACTGGTGAGAGATGATTTTGTATGAGAAATGTCAGCCAGGTCCTCATGAACATGACC 316
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 317 ATCACAGGTGTATGATGACGTTTTCAGGACTACGCGGCTGTGCTTTCGCGAGTGG 376
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 377 ATTAACCAAGCTCTCAATGCCGTGTCATTAACCAACAGAGTGGAGACGCCCTC 436
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 437 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGCACAACACTGGTCCGTAGCAACAGCT 496
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 497 CTAGGACTCAATGCAATTGACCAAGCATGTCTCCACCACATGATAGAGCGTTTGTCCCTTT 556
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db 557 GCTGCCGTAAGCTGCTGCTAATTGCAATTAATTCATTATGAGGCAAGGAACTCAAA 616
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 617 GTTGGCATCCCGTCACGAGATGAGAAATGGAAACCGCTTGGGGAGTCGGCGAAGCGTGG 676
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 677 AAACAAGCCATCACGCAAGTGTGCTGCCAGGATTCATGCGCAGCCCTGGCATGGCC 736
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 737 ATCCCTCGATTCATTATGACACCTTGGAAAAAGAAAGCCCTTTTGAAG 784
```

RESULT 11
AF327346 1818 bp mRNA linear PRI 01-JAN-2002
LOCUS
DEFINITION Homo sapiens tricarboxylate carrier protein mRNA, complete cds.
ACCESSION AF327346
VERSION AF327346.1 GI:18028274
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1818)
AUTHORS Li, N., Chen, T., Wan, T., Zhang, W. and Cao, X.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Department of Immunology, Second Military
Medical University & Shanghai Brilliance Biotechnology Institute,
800 Xiangyin Rd., Shanghai 200433, P.R. China

FEATURES

source

1. 1818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
62. 1030
/codon_start=1
/product="tricarboxylate carrier protein"
/protein_id="AAL56007.1"
/db_xref="GI:18028275"
/translation="MSGELPPNINIKPEPRWDSTFIGRANHEFTVDPNILLTNEQL
ESARKIVHDYRQGIIVPGLTENELMRKVIYDSAFHPDTEKMIIGRMSAQVPMNMT
ITGCMWTFYRTTPAVLFWQWINGSFNVAVNVTNRSGDAPLTVNELGTAVYSATTGAVA
TALGINALTKHVSPLIGRFVFAVAANNCINIPLMRQRELKVGIPVDENGRILGES
ANAAKQAITQVVSRILMAAPGAIPPEIMNTLEKKAFLKRFPMNSAPIQVGLVGFCL
VFATPLCCALFPQKSSMSVTSLEAEIQAKIQESHPELRVYFNKGL"

BASE COUNT 551 a 375 c 403 g 489 t
ORIGIN

Alignment Scores:
Pred. No.: 1.38e-91 Length: 1818
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: Gaps: 9

US-09-990-415A-2 (1-261) x AF327346 (1-1818)

```
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 68 GGAGAACTACCAACCAACATTACATCAAGAACCTCGATGGATCAAGCACTTCATT 127
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
Db 128 GGACGAGCCATCATTTCTCTCACTGTACTGACCCCAAGAACATCTCTTAACCAACGAA 187
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 188 CAACCTCGAGAGTGCAGAAAAATAGTACATGATTACAGGCAAGGAATGTTCTCTGCT 247
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 248 CTTACAGAAATGAATTGTGGAGAGCAAAAGTACATCTATGATTACAGCTTTTCATCCTGAC 307
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 308 ACTGGTGAGAGATGATTTTGTATGAGAAATGTCAGCCAGGTCCTCATGAACATGACC 367
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 368 ATCACAGGTGTATGATGACGTTTTCAGGACTACGCCGCTGTGCTGTCTGCGAGTGG 427
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 428 ATTAACCAAGTCTTCAATGCCGTGCTGTCATTAACCAACAGAAAGTGGAGACGCCCTC 487
```


QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 488 ACTGTCATGAGTTGGGAACAGCTTACGTTTCGACACAACTGTCGCCGTAGACACAGCT 547
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 548 CTAGGACTCAATGCAATGACCAAGCATGTCTCACCACTGATAGACGCTTTGTCCTTT 607
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
Db 608 GCTGCCGTAGCTGCTGCTAATTCATTAATTCATTAATGAGGCAAGGCAACTCAAA 667
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 668 GTTGCATTCCTCCGTCACGGATGAGATGGGACCGCTTGGGGAGTCGGCGAACGCTGCG 727
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 728 AAACAAGCCATCACGCAAGTGTGCTGTCACGAGATTCTCATGTGCAGCCCTGGCATGGCC 787
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 788 ATCCCTCCATTCATTATGACACCTTGGAAAGAAAGCCTTTTGAAG 835
RESULT 12
AK056915 1839 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens CDNA FLJ32353 fls, clone PROST2007612, highly similar
to Tricarboxylate carrier.
ACCESSION AK056915
VERSION AK056915.1 GI:16552444
KEYWORDS oligo capling; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 1839)
REFERENCE Isogai,T., Otsuki,T. and Sugiyama,T.
AUTHORS Direct Submission
TITLE
JOURNAL
COMMENT Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
source
1.1839
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PROST2007612"
/issue_type="prostate"
/clone_lib="PROST2"
/note="cloning vector: pME18SFL3"
BASE COUNT 523 a 399 c 407 g 510 t

ORIGIN
Alignment Scores:
Pred. No.: 1.4e-91 Length: 1839
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 9 Gaps: 0
US-09-990-415A-2 (1-261) x AK056915 (1-1839)
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 63 GGAGAACTACCACCAACAATTACATCAAGGAACCTCGATGGATCAAGCACTTTCATT 122
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
Db 123 GGACGAGCCAAATCATTTCTTCACTGTAACCTGACCCAGGAACATTTCTGTTAACCAACGAA 182
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 183 CAACCTGAGAGTCGAGAAAAAATAGTACATGATTAACAGCAAGAAATGTTCTCTGCT 242
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 243 CTACAGAAATGAATGTTGTGAGAGCAAGATACATCTATGATTCAGCTTTTCATCCTGAC 302
QY 86 ThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
Db 303 ACTGGTGAGAGATGATTTTGTATAGGAAGATGTCAGCCAGGTTCCATGAACATGACC 362
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 363 ATCACAGGTTTATGATGACGTTTACAGACTACGCCGGCTGTGCTGTCTGCGAGTGG 422
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 423 ATTAACAGTCTTCAATGCGCTGCTCAATTACACCAACAGAAAGATGAGATGCACCCCTC 482
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 483 ACTGTCATGAGTTGGGAACAGCTTACGTTCTGCAACAACCTGTCGCTAGCAACAGCT 542
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 543 CTAGGACTCAATGCAATGACCAAGCATGTCTCACCACTGATAGACGTTTGTTCCTTT 602
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
Db 603 GCTGCCGTAGCTGCTGCTAATTCATTAATTCATTAATGAGGCAAGGCAACTCAAA 662
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 663 GTTGCATTCCTCCGTCACGGATGAGATGGGAACCGCTTGGGGAGTCGGCGAACGCTGCG 722
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 723 AAACAAGCCATCACGCAAGTTGTGCTCCAGGATTCTCATGCGACCCCTGGCATGGCC 782
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 783 ATCCCTCCATTCATTATGACACCTTGGAAAGAAAGCCTTTTGAAG 830
RESULT 13
AX302019 2367 bp DNA linear PAT 30-NOV-2001
LOCUS Sequence 18 from Patent WO0177174.
DEFINITION AX302019
ACCESSION AX302019
VERSION AX302019.1 GI:17382991
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

BASE COUNT		684 a	452 c	451 g	649 t
ORIGIN		TKAAQQAIAQVVVSRIGMAAPAMAIIPPVIMNALEKRAFLKRYPMWNPALQVGLVGFCLVFATPLCCALFEPQKSSMSLSRLEPEVLAELKKSPTDKIVYFNKGL"			
Alignment Scores:					
Pred. No.:	2.84e-90	Length:	2236		
Score:	1076.00	Matches:	203		
Percent Similarity:	88.63%	Conservative:	23		
Best Local Similarity:	79.61%	Mismatches:	29		
Query Match:	80.00%	Indels:	0		
DB:	5	Gaps:	0		
US-09-990-415A-2 (1-261) x BC044027 (1-2236)					
QY	7	GluleuProleuAspIleasnIleGlnGluProargTrpaspGlnSerThrPheLeuGly	26		
DB	36	GAACCTGCCCTTCAACATCGACATCAAGAACCACATGGGATCAGAGCACCTTGAGGGA	95		
QY	27	ArgAlaargHisPhePheThrValThrAspProargAsnLeuLeuSerGlyAlaGln	46		
DB	96	AGAGCCAAGCAGCTTCTTACAGTAGACCGATCCACGGAATGCTGCTGGGCCATAAGTC	155		
QY	47	LeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGlyIle	66		
DB	156	CTAGAGGAGCAAGCAACATGTGTACAGAATTACAGAAAGGGGATCCCCACACAGGTTTA	215		
QY	67	ThrGluaspGlnLeuTrpArgAlaAlaLysTyrValTyrAspSeraAlaPheHisProaspThr	86		
DB	216	ACTGAAGACCAGCTATGCGCAGCAAAATACATCTACGATTCTGCTTCCATCCGGATACA	275		
QY	87	GlyGluLysValValLeuIleGlyArgMetSeraGlnValProMetasnMetThrIle	106		
DB	276	GGAGAGAGATGATCTTGATAGGAAGGATGTGCACAGGTGCCCATGAATAGACCATC	335		
QY	107	ThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrpVal	126		
DB	336	ACTGGCTGCATGTGCATTTTTCAGAACATCCAGCTGTTATATTTTGGCAGTGGGTG	395		
QY	127	AsnGlnSerPheAsnAlaIleValAsnTyrSeraAsnArgSergLysPthrProIleThr	146		
DB	396	AATCAGTCCCTTCAATGCTATTGTGAATTATACAAATAGAAGTGTGATGACCCATAACA	455		
QY	147	ValArgGlnLeuGlyThrAlaTyrValSeraIleThrThrGlyAlaValAlaThrAlaLeu	166		
DB	456	ACAAGCCAGCTGGGAGACAGCTTACGTACGTGCACTACAGGTGCTGTGTGACTGCTCTG	515		
QY	167	GlyLeuLysSerLeuThrLysHisLeuProProleuValGlyArgPheValProPheAla	186		
DB	516	GGACTCAAAATCACTCACTAAGCATCTACCTTCTGTGGTGGAGATTGTTCATTGGCC	575		
QY	187	AlaValAlaAlaAlaAsnCysIleAsnIleProleuMetArgGlnArgGluLeuGlnVal	206		
DB	576	GCAGTAGACAGCAGCAACTGCATCAACATCCCTCTGATGAGGCAGAGAGAAATTAATAATAT	635		
QY	207	GlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSeraValThrAlaAlaLys	226		
DB	636	GGCATCCCATCACAGATGAAAATGGCAACAGGCTCGGAGAAATCTACCAAGCTGCTCAG	695		
QY	227	GlnGlyIlePheGlnValValIleSeraGlyIleCysMetAlaIleProAlaMetAlaIle	246		
DB	696	CAGGCCATAGCACACAAGTGTGTCTCCCGATTGGAATGGCTGCCCCAGCAATGGCAATT	755		
QY	247	ProProLeuIleMetaspThrLeuGluLysLysaspPheLeuLys	261		
DB	756	CCCCCAGTATATGAATGCTTTGGAGAAAGACGCTTCTCTAAAG	800		
RESULT 15					
S70011	S70011	2986 bp	mRNA	linear	ROD 10-JUL-2000
LOCUS					
DEFINITION					
ACCESSION					
gene for mitochondrial product.					

VERSION	S70011.1	GI:545997
KEYWORDS		
SOURCE	Rattus sp.	
ORGANISM	Rattus sp.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 2986)	
AUTHORS	Azzl,A., Glerum,M., Koller,R., Mertens,W. and Spycher,S.	
TITLE	The mitochondrial tricarboxylate carrier	
JOURNAL	J. Bioenerg. Biomembr. 25 (5), 515-524 (1993)	
MEDLINE	94179133	
PUBMED	8132491	
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI g1bbsq 147646] from the original journal article. This sequence comes from Fig. 9.	
FEATURES	location/Qualifiers	
source	1..2986	
	/organism="Rattus sp."	
	/mol_type="mRNA"	
	/db_xref="taxon:10118"	
	/tissue_type="liver"	
	<1..1076	
	/note="This sequence comes from Fig. 9"	
	/codon_start=3	
	/product="tricarboxylate carrier"	
	/protein_id="AAB30258.1"	
	/db_xref="GI:545998"	
	/translation="APGGGARDRGASGSGRRHGAGPRGAGGGRGGSGTMSGEVPPNI NIKEPRWDSTFIGRASHFTYTDPKNILLTNEOLENARKVHDYRGTVPAGLTENE LMRAKYAYDSAFHPDTEKMTLIGRMSAOVPMNMTITGCMTFYRTTPAVLFWQWINO SFNAVNVNTRSGDAPLVNELGTAYVSAFTGAVATAGLNATKHSPLIGRFVFA AVAAANCINIPLMRQRELKVGIPVTDENGTRLGESTNAAKOAITOVVISRIIMAAPGM AIPFIMNTLEKKAFLKRFPMWSAPIQVTLVGLVFATPLCCALFFQKSSMSVTSLE DDLQASIQSHPELRVRYFNKGL"	
BASE COUNT	818 a	690 c 697 g 781 t
ORIGIN		
Alignment Scores:		
Pred. No.:	5.22e-89	Length: 2986
Score:	1064.00	Matches: 198
Percent Similarity:	87.89%	Conservative: 27
Best Local Similarity:	77.34%	Mismatches: 31
Query Match:	79.11%	Indels: 0
DB:	10	Gaps: 0
US-09-990-415A-2 (1-261) x S70011 (1-2986)		
QY	6	GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
DB	114	GGGGAAGTGCACCCCAACATTAACATCAAGAGCCTCGATGGACCAGACAGCTTCATT 173
QY	26	GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
DB	174	GGCCGAGCCAGTCACTTCTTACCGGTTACTGATCCCAAAACATCCTTTTAACGAACGAA 233
QY	46	GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
DB	234	CAGCTAGAGAAATGCGAAGAAAGTGTACACGATTAACAGCAAGGAATCGTTCTCGCGGC 293
QY	66	IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSeraAlaPheHisProAsp 85
DB	294	CTCACGGAATAATGAGTTATGGAGAGCGAAGTACCGGTATGACTCGGCTTCATCCCGAC 353
QY	86	ThrGlyGluLysValValLeuIleGlyArgMetSeraAlaGlnValProMetAsnMetThr 105
DB	354	ACGGGTGAATAAGATGACTGTGATAGGAAGAATGTGGCTCAGTCCCGATGAACATGACC 413
QY	106	IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
DB	414	ATCACGGGCTGCATGATGACCTTCTACCGGACACACACCGGCTGTGCTTTCTGCGAGTGG 473
QY	126	ValAsnGlnSerPheAsnAlaIleValAsnTyrSeraAsnArgSergLysPthrProIle 145

Db 474 ATAAACCAAGTCCTTCATGCGCGTGCACACTACACCAATAGAGAGTGGTGACGCTCCCTT 533
QY 146 ThrValArgInLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 534 ACTGTAATGAGCTGGGAACGGCTTATGTCTGCCAACACCGGCCGTGGCGACAGCC 593
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValProPhe 185
Db 594 CTGGACTCAATGCCCCTAACCAAGCATGTGCGCCGCTGATAGACGTTGTCCCTTC 653
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db 654 GCTGCTAGCTGCTGCTAACTGATTAACATCCCGCTGATGAGACAAAGGAACTCAAG 713
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 714 GTTGGTATTCCTGTCACTGATGAGAATGGCACACGCTTGGAGAAATCGAACGACAGCA 773
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 774 AAGCAAGCCATCACGCAAGGTGTCATCTCCAGGATCCTCATGGCGGCCCTGGCATGGCC 833
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 834 ATCCCTCCGTTTATCATGAACACCTTGGAGAAAGCCCTTTTGAAG 881

Search completed: October 16, 2003, 16:59:43
Job time : 2923 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 13:25:31 ; Search time 101 Seconds
(without alignments)
5384.000 Million cell updates/sec

Title: US-09-990-415A-1
Perfect score: 1232
Sequence: 1 cccttagcgccagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94.4	76.6	1729	4 US-09-620-312D-1016	Sequence 1016, Ap
2	65.8	5.3	315	3 US-08-905-223-165	Sequence 165, App
3	41.6	3.4	2303	4 US-09-922-146-3	Sequence 3, Appl1
4	41.2	3.3	1894	3 US-07-728-220C-1	Sequence 1, Appl1
5	40.8	3.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
6	40.8	3.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
7	39.8	3.2	642	4 US-09-252-991A-7687	Sequence 7687, Ap
8	39.8	3.2	1761	4 US-09-252-991A-7943	Sequence 7943, Ap
9	39.8	3.2	2718	4 US-09-252-991A-7538	Sequence 7538, Ap
10	39.8	3.2	3351	4 US-09-252-991A-7803	Sequence 7803, Ap
11	38.4	3.1	6453	1 US-08-306-691B-14	Sequence 14, Appl
12	38.4	3.1	6453	3 US-09-209-668-10	Sequence 10, Appl
13	38.4	3.1	6453	3 US-09-356-952-8	Sequence 8, Appl1
14	38.2	3.1	816	4 US-09-252-991A-11434	Sequence 11434, A
15	38.2	3.1	2028	4 US-09-252-991A-11416	Sequence 11416, A
16	37.6	3.1	822	4 US-09-252-991A-11456	Sequence 11456, A
17	37.2	3.0	243	3 US-09-461-697-463	Sequence 463, App
18	37.2	3.0	979	3 US-09-461-697-462	Sequence 462, App
19	37	3.0	3526	4 US-09-620-312D-182	Sequence 182, App
20	37	3.0	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
21	37	3.0	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
22	36.6	3.0	927	4 US-09-252-991A-1510	Sequence 1510, Ap
23	36.6	3.0	1893	4 US-09-252-991A-1462	Sequence 1462, Ap
24	35.8	2.9	1767	4 US-09-016-434-1488	Sequence 1488, Ap
25	35.8	2.9	15378	3 US-08-785-420-1	Sequence 1, Appl1
26	35.6	2.9	1152	4 US-09-252-991A-7652	Sequence 7652, Ap
27	35.6	2.9	1383	4 US-09-252-991A-7737	Sequence 7737, Ap

C	28	35.6	2.9	2997	4 US-09-252-991A-7831	Sequence 7831, Ap
	29	35.4	2.9	1308	3 US-09-188-930-248	Sequence 248, App
	30	35.4	2.9	1308	4 US-09-312-283C-248	Sequence 248, App
	31	35.4	2.9	1421	3 US-09-188-930-244	Sequence 244, App
	32	35.4	2.9	1421	4 US-09-312-283C-244	Sequence 244, App
C	33	35.2	2.9	789	4 US-09-252-991A-2208	Sequence 2208, Ap
	34	35.2	2.9	1095	4 US-09-252-991A-2466	Sequence 2466, Ap
	35	35	2.8	2150	2 US-08-318-837-1	Sequence 1, Appl1
C	36	35	2.8	2897	2 US-08-927-394-1	Sequence 1, Appl1
	37	35	2.8	2897	4 US-09-016-434-1163	Sequence 1163, Ap
C	38	34.8	2.8	435	4 US-09-395-674B-5	Sequence 475, App
	39	34.8	2.8	1133	4 US-09-252-991A-475	Sequence 5, Appl1
C	40	34.8	2.8	1569	4 US-09-395-674B-5	Sequence 423, App
	41	34.8	2.8	1617	4 US-09-252-991A-381	Sequence 381, App
C	42	34.8	2.8	1683	4 US-09-252-991A-5772	Sequence 5772, Ap
	43	34.8	2.8	1722	4 US-09-252-991A-5811	Sequence 5811, Ap
C	44	34.8	2.8	2186	2 US-08-878-546-9	Sequence 9, Appl1
	45	34.8	2.8	2772	4 US-09-252-991A-5838	Sequence 5838, Ap

ALIGNMENTS

RESULT 1
US-09-620-312D-1016
; Sequence 1016, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrul
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620, 312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (247)..(1371)
US-09-620-312D-1016

Query Match 76.6%; Score 94.4; DB 4; Length 1729;
Best Local Similarity 100.0%; Pred. No. 9.3e-248;
Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 289 GGTTCGCAATCCCGTGGCCACCGGCGGGCGGGGAGAGCTCTCCCTCCCT .348
Db 86 GGTTCGCAATCCCGTGGCCACCGGCGGGCGGGGAGAGCTCTCCCTCCCT 145


```

QY 349 GCTGTCGGCGCTACCGCTGACGTCCCGCGTGATGGCTGGAGGGCCCCGGGGCGACAGC 408
Db 146 GCTGTCGGCGCTACCGCTGACGTCCCGCGTGATGGCTGGAGGGCCCCGGGGCGACAGC 205
QY 409 GGAGGCAGAGAGGAGGGCGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 468
Db 206 GGAGGCAGAGAGGAGGGCGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 265
QY 469 AATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGAGCCAAAGTACTTTCCTGGGA 528
Db 266 AATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGAGCCAAAGTACTTTCCTGGGA 325
QY 529 GAGCCCGGCGACTTTTCACTGTACTGATCCTCGAAATCTGCTGTCTCCGGGGCACAGC 588
Db 326 GAGCCCGGCGACTTTTCACTGTACTGATCCTCGAAATCTGCTGTCTCCGGGGCACAGC 385
QY 589 TGGAGCTTCTCGGACATCTGCGAAGACTACAGAGGGCGCGGTGTGACCCAGGGATCA 648
Db 386 TGGAGCTTCTCGGACATCTGCGAAGACTACAGAGGGCGCGGTGTGACCCAGGGATCA 445
QY 649 CCGAGGACCGAGCTGTGGAGGGCGCAAGTATGTATGATCCGCTTCATCCGGACAG 708
Db 446 CCGAGGACCGAGCTGTGGAGGGCGCAAGTATGTATGATCCGCTTCATCCGGACAG 505
QY 709 GGGAGAGGCTGTGCTGATGGCGCGCATGTGAGCCCGAGTCCCATGAACATGACCATCA 768
Db 506 GGGAGAGGCTGTGCTGATGGCGCGCATGTGAGCCCGAGTCCCATGAACATGACCATCA 565
QY 769 CTGGCTGATGCTCACATTTCTACAGGAAGACCCCAACCGTGTGTCTTGCGAGTGGTGA 828
Db 566 CTGGCTGATGCTCACATTTCTACAGGAAGACCCCAACCGTGTGTCTTGCGAGTGGTGA 625
QY 829 ATCAGTCTTCAATGCCATTTTAATACTACTCCAACCGGAGTGTGACACTCCCATCACTG 888
Db 626 ATCAGTCTTCAATGCCATTTTAATACTACTCCAACCGGAGTGTGACACTCCCATCACTG 685
QY 889 TGAAGCAGCTGGGGACAGCCCTATGTGAGTGCCACCACTGAGAGCTGTGGCCAGGCCCTGG 948
Db 686 TGAAGCAGCTGGGGACAGCCCTATGTGAGTGCCACCACTGAGAGCTGTGGCCAGGCCCTGG 745
QY 949 GACTCAAAATCCCTCACCAGCACTGCCCCCTGTGGCGAGATTGTGCCCTTTGCGAG 1008
Db 746 GACTCAAAATCCCTCACCAGCACTGCCCCCTGTGGCGAGATTGTGCCCTTTGCGAG 805
QY 1009 CAGTGGCAGCTGCCAATCTGATCAACATCCCCCTGATGAGCAGAGAGAGCTGACAGTGG 1068
Db 806 CAGTGGCAGCTGCCAATCTGATCAACATCCCCCTGATGAGCAGAGAGAGCTGACAGTGG 865
QY 1069 GCATCCCGGTGGCTGATGAGGCGAGGCTGAGAGGCTTGGCTACTCGGTGACTGCAGCAAGC 1128
Db 866 GCATCCCGGTGGCTGATGAGGCGAGGCTGAGAGGCTTGGCTACTCGGTGACTGCAGCAAGC 925
QY 1129 AGGGAATCTTCCAGGTGTGATTTCAAGATCTGCATGGCGATTCCTGCCATGGCCATCC 1188
Db 926 AGGGAATCTTCCAGGTGTGATTTCAAGATCTGCATGGCGATTCCTGCCATGGCCATCC 985
QY 1189 CACCACGTGATGATGACACTCTGAGAGAAAGAGACTTCTGAAG 1232
Db 986 CACCACGTGATGATGACACTCTGAGAGAAAGAGACTTCTGAAG 1029

```

```

RESULT 2
US-08-905-223-165
; Sequence 165, Application us/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duelet, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear

```

```

; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: us/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 165:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 70..108
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 3.5
; OTHER INFORMATION: seq MHLISNMWNPASS/RR
; US-08-905-223-165

```

```

Query Match 5.3%; Score 65.8; DB 3; Length 315;
Best Local Similarity 56.1%; Pred. No. 1.2e-08;
Matches 124; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 496 AACCTGCTGGGACCAAGTACTTTCCTGGGACAGAGCCCGCACTTTTCACTGTTACTG 555
Db 94 AACCCCGCTTCAGACAGACGCTCTTATGCGCGCTTCAGGCACTTCTTGATATCATCG 153
QY 556 ATCTCGAAATCTGCTGCTGTCGGGGGACAGCTGGAAGCTTCTCGAATCGTGAGA 615
Db 154 ACCCTCGACACTCTTGTCTAGAGAGCGTCTCAGAGAGGCTGTGACAGCTGTGAGG 213
QY 616 ACTACAGGGCGCGTGTGTAACCCAGGATCACCGAGACCAAGCTGTGAGGGCCAGT 675
Db 214 ACTATAAGCATGGGACCCCTGGCGGGGGGTACCAATGAACAGCTCTGAGTGCACAGA 273
QY 676 ATGTGTATGACTCCGCTTCATCCGGACACAGGGGAGAAG 716
Db 274 AATCAAGCAGGCTATTCTACATCCGGACACCAATGAGAAG 314

```

```

RESULT 3
US-09-922-146-3
; Sequence 3, Application us/09922146
; Patent No. 6566133
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
; FILE REFERENCE: RTS-0252
; CURRENT APPLICATION NUMBER: US/09/922,146
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3

```

```

; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(1268)
US-09-922-146-3

```

```

Query Match          3.4%; Score 41.6; DB 4; Length 2303;
Best Local Similarity 51.3%; Pred. No. 0.1;
Matches 122; Conservative 0; Mismatches 114; Indels 2; Gaps 1;

```

```

QY 200 GCACCCCTTGAGGCTCCCTGAGAGGAGCAGTCACGGGGAGAGCGGAGCGTAGAGCTGAGT 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 GCAGCCGCGAGCTGTACGAGTCGCGCGCGCATCGGTGGGCGCTGAGCGTGCGCCCTGCCG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 260 TGCTGCAG--GCGTGTCTGTGTGTGTGGCTGGGTTCGCCAATCCCCGTGCCACCGGGT 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CGCTCCTGCTGCGCGCGCGCGGAGGGGACGCTGCGGTGCCGCGCGCTCCTGCCCTGGC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 318 GGGCGCGCGCGGAGCTCCTGCCCTCCTCCTGCTGCTGCGCGTCACGCGTAGACGTCGCC 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 CGCCGCTGACGCGCGCGCGCGCGCTGCGCGCTGCTGCTGACGACGAGGCGGGGCGCGGC 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 378 GTGATGCTGGAGGCGCGCGCGCGGACGACGCGGAGGAGGAGGAGGCGGTCTGAG 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 GCGCGCGCGGAGGCGCGGAGGCGCGGAGGCGGAGGAGGAGGCGGTGCTGCTGG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

```

US-07-728-220C-1
; Sequence 1, Application US/07728220C
; Patent No. 6020168

```

```

; GENERAL INFORMATION:
; APPLICANT: MATSUO, HISAYUKI
; APPLICANT: KANGAWA, KENJI
; APPLICANT: TANAKA, SHOJI
; APPLICANT: FUCHIMURA, KAYOKO
; APPLICANT: TAWARAGI, YASUNORI
; TITLE OF INVENTION: PORCINE CNP GENE AND PRECURSOR PROTEIN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: ELEVENTH FLOOR, 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,220C
; FILING DATE: 19910712
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/9437/91816
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; TELEX: 671 4627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1894 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-728-220C-1

```

```

Query Match          3.3%; Score 41.2; DB 3; Length 1894;
Best Local Similarity 59.3%; Pred. No. 0.12;
Matches 70; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

```

QY 305 GTGCCACCGGGGTGGGCGCGCGGGAAGCTCCTGCCCCCTGCTGCTGCGGCTCAGC 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14 GTGGGAAGAGGGTGGGGGCGAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 CGTACGCTCCCGCTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 CGTGCCATTGGCGCGCGCGCGCTTGTGGGCGGAGGATGATCATCAGCGGAGGTTGA 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

```

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328

```

```

; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

```

Query Match          3.3%; Score 40.8; DB 3; Length 4403765;
Best Local Similarity 54.7%; Pred. No. 3.7;
Matches 81; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```

```

QY 302 CCCGTGCCACCGGGGTGGGCGCGCGGGAAGCTCCTGCCCCCTGCTGCTGCGGCTC 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2210214 CCCGCGACGACGACGAGGCGGCTGCGCAACACGAGGCGGCGAGCTGTGAGTTGCTGTC 2210155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 ACGGTGACGTCGCCGCTGATGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2210154 ACGGTGCTCTCCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2210095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 AAGCGGTTCTGAGAGCTTCAGAGCG 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2210094 CGGGCGGCTCGCGCGGCTACGATACCG 2210067
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328

```

```

; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1

```

; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match 3.3%; Score 40.8; DB 3; Length 4411529;
Best Local Similarity 54.7%; Pred. No. 3.7;
Matches 81; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 302 CCCGTGCCACCGGGTGGCGCGCGGGAAGCTCCTGCCCTCCTGCTGCTGCGCGTC 361
Db 2212914 CCCGCGACACGACACGAGGGGCTCGCCACACCGAGCGGCGACGCTGAGTTTCGTTGTC 2212855
QY 362 ACGCGTGACGTCCCGCGTGAATGCTGGAGGGCCCCGCGCGACAGCGGAGCGACAGAG 421
Db 2212854 ACGGTGTGCTCTCCCGCTGCTGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCG 2212795
QY 422 AAGCGGCTTCTGAGAGCTTCAGAGACG 449
Db 2212794 CGGGCGGCTCCGGCGGCTACGATAGCG 2212767

RESULT 7
US-09-252-991A-7687/c
; Sequence 7687; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7687
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7687

Query Match 3.2%; Score 39.8; DB 4; Length 642;
Best Local Similarity 48.5%; Pred. No. 0.19;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 779 GCTCACATTTCTACAGAGACCCCAACCGTGTCTGTCAGTGGTGAATCAGTCTT 838
Db 584 GCTGGGCTTCGCCAAGAGCTGGCGCGGAGGAGAGACTCCAGCGCGGAGGCGCTGT 525
QY 839 CAATGCCATTTGTAATCTCCCAACCGGAGTGTGACACTCCCATCACTGTGAGGACGCT 898
Db 524 CCGTGGCATGTTGACAGAGCCCTACCGCAACGGGCAAGCGGCTGAACATGGCCACTACCT 465
QY 899 GGGACAGCCCTATGTGAGTGGCCACCACTGAGAGCTGTGGCCACGGCCCTGGAGTCAATC 958
Db 464 GGAGATCGACGGCGGTGATGACCCCGGCGGAGAGACCCGCGCTGCTCCTGCGCGGCTGGC 405
QY 959 CCTCACCAGCACCCTGCCCCCTGTGTCGGGACAGATTGTGCCCTTG 1005
Db 404 CGTCGCCGAGAACCGCGCGGAGGGCGGCGCAAGCGGCGCTTCG 358

RESULT 8
US-09-252-991A-7943
; Sequence 7943; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7943
; LENGTH: 1761
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7943

Query Match 3.2%; Score 39.8; DB 4; Length 1761;
Best Local Similarity 48.5%; Pred. No. 0.29;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 779 GCTCACATTTCTACAGAGACCCCAACCGTGTCTGTCAGTGGTGAATCAGTCTT 838
Db 1360 GCTGGGCTTCGCCAAGAGAGCTGGCGCGGAGGAGAGACTCCAGCGGCGGAGCGGCTGT 1419
QY 839 CAATGCCATTTGTAATCTCCCAACCGGAGTGTGACACTCCCATCACTGTGAGGACGCT 898
Db 1420 CCGTGGCATGTTGACAGAGCCCTACCGCAACGGGCAAGCGGCTGAACATGGCCACTACCT 1479
QY 899 GGGACAGCCCTATGTGAGTGGCCACCACTGAGAGCTGTGGCCACGGCCCTGGAGTCAATC 958
Db 1480 GGAGATCGACGGCGGTGATGACCCCGGCGGAGAGACCCCGCTGCTCCTGCGCGGCTGGC 1539
QY 959 CCTCACCAGCACCCTGCCCCCTGTGTCGGGACAGATTGTGCCCTTG 1005
Db 1540 CGTCGCCGAGAACCGCGCGGAGGGCGGCGCAAGCGGCGCTTCG 1586

RESULT 9
US-09-252-991A-7538/c
; Sequence 7538; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7538
; LENGTH: 2718
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7538

Query Match 3.2%; Score 39.8; DB 4; Length 2718;
Best Local Similarity 48.5%; Pred. No. 0.34;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 779 GCTCACATTTCTACAGAGACCCCAACCGTGTCTGTCAGTGGTGAATCAGTCTT 838
Db 268 GCTGGGCTTCGCCAAGAGAGCTGGCGCGGAGGAGAGACTCCAGCGGCGGAGGCGCTGT 209
QY 839 CAATGCCATTTGTAATCTCCCAACCGGAGTGTGACACTCCCATCACTGTGAGGACGCT 898
Db 208 CCGTGGCATGTTGACAGAGCCCTACCGCAACGGGCAAGCGGCTGAACATGGCCACTACCT 149
QY 899 GGGACAGCCCTATGTGAGTGGCCACCACTGAGAGCTGTGGCCACGGCCCTGGAGTCAATC 958
Db 899 GGGACAGCCCTATGTGAGTGGCCACCACTGAGAGCTGTGGCCACGGCCCTGGAGTCAATC 958

Db 148 GGAGATCGACGGGTGATCGACCCGGCGAGACCCGCCCTGGCTCCTGCGCGGCTGGC 89
QY 959 CCTCACCAAGCAGCTGCCCCCTGGTCGGCAGATTGTGCCCTTTG 1005
Db 88 CGTCGCCGGAGAACCGCGCCGAGGGCGGCGCAAGCGGCCGTTG 42

RESULT 10

US-09-252-991A-7803
; Sequence 7803, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7803
; LENGTH: 3351
; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7803

Query Match 3.2%; Score 39.8; DB 4; Length 3351;
Best Local Similarity 48.5%; Pred. No. 0.37;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 779 GCTCATATTCTACAGGAGACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCCCT 838
Db 3111 GCTGGGCTTCGCCAAGAGAGCTGGCGCGCGAGGAAGACTCCAGCGGCGAGGCGCTGTT 3170
QY 839 CAATGCCATTGTTACTACTCCCAACCGCAGTGTGACACTCCCATCACTGTGAGGAGCT 898
Db 3171 CCGTGGCATGTGTGACCAAGGCTTACCGCAACGCGCTGAACATGCGCAGCTACCT 3230
QY 899 GGGAGACGCTATGTAGTGCACACCTGAGAGCTGTGGCCACGCGCTGGGAGCTCAATC 958
Db 3231 GGAGATGACGCGGTGATCGACCCGCGGAGAGACCCGCTGCTGCTGCGGCGCTGGC 3290
QY 959 CCTCACCAAGCAGCTGCCCCCTGGTCGGCAGATTGTGCCCTTTG 1005
Db 3291 CGTCGCCGGAGAACCGCGCCGAGGGCGGCGCAAGCGGCCGTTG 3337

RESULT 11

US-08-306-691B-14
; Sequence 14, Application US/08306691B
; Patent No. 5734039

GENERAL INFORMATION:

; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgha & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: NO. 5734039e

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 6453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-306-691B-14

Query Match 3.1%; Score 38.4; DB 1; Length 6453;
Best Local Similarity 44.8%; Pred. No. 1.2;
Matches 147; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 87 GCTCATTCGCAATTAGGCTTGGCTGTGCTGTTCCTCCCTCGGCTGAACCACTC 146
Db 576 GGTAGTGGCGCAGGATCGCGCGCGCCGCCCTCTCGCCCAAGCGGAGCAA 635
QY 147 TTCTGTAGCCGAGCAGCTACCGGGGCTGTGGAATTGCCACCCCTCCCTGGGACCC 206
Db 636 TACG 695
QY 207 TGAGGCTCCGTGAGGAGGAGCTACGCGGCGGAGAGCGGAGCTGAGCTGATTGCTGCA 266
Db 696 GTGGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 755
QY 267 GCGGTCTCTGT 326
Db 756 GCGGGGGGCGCTTCT 815
QY 327 CGGGAAGCTCTGCCCCCTCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
Db 816 CTGGGACAGCTCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 875
QY 387 GGGAGGCG 414
Db 876 GGGGCTGTGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903

RESULT 12

US-09-209-668-10
; Sequence 10, Application US/09209668A
; Patent No. 6114517

GENERAL INFORMATION:

; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: ALPHA-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0336
; CURRENT APPLICATION NUMBER: US/09/209,668A
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1664)..(1774)
; FEATURE:

NAME/KEY: CDS
LOCATION: (2042)..(2220)
FEATURE:
NAME/KEY: CDS
LOCATION: (2374)..(2533)
FEATURE:
NAME/KEY: CDS
LOCATION: (3231)..(3350)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: V00574/Genbank
DATABASE ENTRY DATE: 1991-01-03
US-09-209-668-10

Query Match 3.1%; Score 38.4; DB 3; length 6453;
Best Local Similarity 44.8%; Pred. No. 1.2;

Matches 147; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 87 GCTCATTCGCAATTAGGCTTGGCTGTGCTTGTCCCTCCGGCTCGAACCACTC 146
Db 576 GGTAGTGCAGGAGGATCGCGCGCGCGCGCGCTCTCCCTCCAGGCGGCA 635
QY 147 TTCTCTGAGCCGAGCCAGCTACCGGGGCTCTGGAATTGCCACCCCTCCCTGGCACTC 206
Db 636 TACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAAGCGCGCGCGCGCG 695
QY 207 TGAGGCTCTCCGTGAGGAGGAGCTACGGGCGAGAGCGGAGCTGAGCTGATTGCTCA 266
Db 696 GTGGGTGGGGCGCGCGCGCGCGCGCGCGCGCGAGTGAAGCGCGCGCTCGCGCG 755
QY 267 GCGTGTCTGTGTGTGGCTGGCTTCTGCCAATCCCGTCCACCGGCTGGCGCGCG 326
Db 756 GCGCGGGGCGCGCTCTCTCTCTGGGGCTGTGCGGAATCCGGCGCGCGCGCTGCGG 815
QY 327 CGGGAAGCTCTCTGCGCTGCTGTGCTGCGCTACGCGGTGAGCTCCCGCTGATGCT 386
Db 816 CTGGGACAGCTCCCGACGCGCGCGCTACCGGAGGCTCGGGCGCGCGCTCACACCG 875
QY 387 GGGAGGGCGCGCGCGCGCGAGCGGAGGC 414
Db 876 GGGGCGTCTGGAGAGGCGCGCGCGCG 903

RESULT 13
US-09-356-952-8

; Sequence 8, Application US/09356952

; Patent No. 6117663

; GENERAL INFORMATION:

; APPLICANT: Boriack-Stodin, Ann

; APPLICANT: Margalit, S. M.

; APPLICANT: Bor-Sogil, Dafna

; APPLICANT: Kuriyan, John

; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE

; FILE REFERENCE: 600-1-228N

; CURRENT APPLICATION NUMBER: US/09/356,952

; EARLIER APPLICATION NUMBER: 60/093,631

; EARLIER FILING DATE: 1998-07-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 6453

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-356-952-8

Query Match 3.1%; Score 38.4; DB 3; length 6453;
Best Local Similarity 44.8%; Pred. No. 1.2;

Matches 147; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 87 GCTCATTCGCAATTAGGCTTGGCTGTGCTTGTCCCTCCGGCTCGAACCACTC 146
Db 576 GGTAGTGCAGGAGGATCGCGCGCGCGCGCGCTCTCCCTCCAGGCGGCA 635

Db 576 GGTAGTGCAGGAGGATCGCGCGCGCGCGCGCGCTCTCTCCAGGCGGCA 635
QY 147 TTCTCTGAGCCGAGCCAGCTACCGGGGCTCTGGAATTGCCACCCCTCCCTGGCACTC 206
Db 636 TACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAAGCGCGCGCGCGCG 695
QY 207 TGAGGCTCTCCGTGAGGAGGAGCTACGGGCGAGAGCGGAGCTGAGCTGATTGCTCA 266
Db 696 GTGGGTGGGGCGCGCGCGCGCGCGCGCGCGCGAGTGAAGCGCGCGCTCGCGCG 755
QY 267 GCGTGTCTGTGTGTGGCTGGCTTCTGCCAATCCCGTCCACCGGCTGGCGCGCG 326
Db 756 GCGCGGGGCGCGCTCTCTCTCTGGGGCTGTGCGGAATCCGGCGCGCGCTGCGG 815
QY 327 CGGGAAGCTCTCTGCGCTGCTGTGCTGCGCTACGCGGTGAGCTCCCGCTGATGCT 386
Db 816 CTGGGACAGCTCCCGACGCGCGCGCTACCGGAGGCTCGGGCGCGCGCTCACACCG 875
QY 387 GGGAGGGCGCGCGCGCGAGCGGAGGC 414
Db 876 GGGGCGTCTGGAGAGGCGCGCGCGCG 903

RESULT 14

US-09-252-991A-11434

; Sequence 11434, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11434

; LENGTH: 816

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-11434

Query Match 3.1%; Score 38.2; DB 4; length 816;
Best Local Similarity 63.7%; Pred. No. 0.57;

Matches 58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 160 GCCAGCTACCGGGGCTCTGGAATTGCCACCCCTCTGGGACACCTTGAGGCTTCGCTG 219
Db 285 GCGCACTACGGGCGCGCTGTGTGCTACGAGCTCCCTGGGACACCGGCGCTCCGAGGG 344
QY 220 GAGGAGCTCACGGGCGAGAGCGGAGCTGA 250
Db 345 GCGAAGCTACGGGCGCGCTCGAGCGTCA 375

RESULT 15

US-09-252-991A-11416

; Sequence 11416, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR APPLICATION NUMBER: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

Fri Oct 17 07:50:57 2003

us-09-990-415a-1.rni

Page 7

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11416
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11416

```

Query Match	3.1%;	Score 38.2;	DB 4;	Length 2028;
Best Local Similarity	63.7%;	Pred. No. 0.83;		
Matches	58;	Conservative	0;	Mismatches 33;
			Indels	0;
			Gaps	0;

QY 160 GCCAGCTACCGGGGCTCTGGAAATTGCCACCCCTCCCTGGGCACCCCTTGAGGCCCTCCGTG 219
 Db 37 GCGGACTACGGGGGCGCGCTGTGTGTGTCTACGGCTCCCTGGGCACCCGCGCGTCCCGAGGG 96
 QY 220 GAGGAGCTCACGGGCGACAGCGGAGCTGA 250
 Db 97 GCGGACGCTACGGGGCGGTCGGACGGTCA 127

Search completed: October 16, 2003, 16:00:09
Job time : 118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 11:00:46 ; Search time 4526 Seconds
(without alignments)
11135.806 Million cell updates/sec

Title: US-09-990-415A-1
Perfect score: 1232
Sequence: 1 cccttagcgccagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1231	99.9	1826	6	AX061199	AX061199 Sequence
2	1208.2	98.1	3047	6	AX188304	AX188304 Sequence
3	848.8	68.9	2702	6	BD127742	BD127742 Primer fo
4	848.8	68.9	2702	9	AK074707	AK074707 Homo sapi
5	845	68.6	2735	9	BC000124	BC000124 Homo sapi
6	818.6	66.4	984	6	AX364964	AX364964 Sequence
7	674	54.7	2870	10	BC012208	BC012208 Mus muscu
8	664.2	53.9	2699	10	AF276997	AF276997 Rattus no
9	659	53.5	752	6	BD124792	BD124792 Primer fo
10	659	53.5	752	6	BD126228	BD126228 Primer fo
11	655.6	53.2	2710	10	AF325262	AF325262 Mus muscu
12	432	35.1	1204	10	BC005743	BC005743 Mus muscu
13	432	35.1	2803	10	AF325260	AF325260 Mus muscu
14	422.4	34.3	2986	10	S70011	S70011 Rattus sp.
15	417.8	33.9	2236	5	BC044027	BC044027 Xenopus 1
16	398.2	32.3	1818	9	AF327346	AF327346 Homo sapi
17	398.2	32.3	2367	6	AX302019	AX302019 Sequence
18	396.8	32.2	1134	6	AX511171	AX511171 Sequence
19	396.6	32.2	1839	9	AK056915	AK056915 Homo sapi
20	366.4	29.7	2894	9	BC020517	BC020517 Homo sapi
21	351.2	28.5	2766	6	BD157602	BD157602 Primer fo
22	351.2	28.5	2766	9	AK022287	AK022287 Homo sapi
23	346.4	28.1	1084	6	AX457932	AX457932 Sequence
24	332.2	27.0	1772	6	AX392949	AX392949 Sequence
25	332.2	27.0	2499	9	AF462052	AF462052 Homo sapi
26	332.2	27.0	2608	9	BC022091	BC022091 Homo sapi
27	315	25.6	984	6	AX457933	AX457933 Sequence
28	315	25.6	984	6	AX511173	AX511173 Sequence
29	305.6	24.8	1537	10	AF325261	AF325261 Mus muscu
30	305.6	24.8	1617	10	BC019808	BC019808 Mus muscu
31	302.8	24.6	2642	6	BD159719	BD159719 Primer fo
32	302.8	24.6	2642	9	AK022938	AK022938 Homo sapi
33	302.4	24.5	1575	6	AX740470	AX740470 Sequence
34	299.6	24.3	2062	5	BC048044	BC048044 Dario rer
35	291	23.6	181086	9	AL133215	AL133215 Human DNA
36	288.4	22.4	741	6	BD148081	BD148081 Primer fo
37	278.6	22.6	1316	3	AY089419	AY089419 Drosophi1
38	278.6	22.6	1394	3	BT001832	BT001832 Drosophi1
39	226.2	18.4	2078	3	AY071029	AY071029 Drosophi1
40	225.6	18.3	524	6	AX552273	AX552273 Sequence
41	222.8	18.1	806	6	BD146776	BD146776 Primer fo
42	204	16.6	2324	6	AX713842	AX713842 Sequence
43	204	16.6	2324	9	AK055711	AK055711 Homo sapi
44	173.6	14.1	76187	2	AC015219	AC015219 Drosophi1
45	173.6	14.1	175963	3	AC009379	AC009379 Drosophi1

ALIGNMENTS

RESULT 1
AX061199
LOCUS AX061199 1826 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 46 from Patent WO0078953.
ACCESSION AX061199
VERSION AX061199.1 GI:12406335
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Lal,P., Yang,J., Yue,H., Hillman,J.L., Tang,Y.T., Bandman,O.,
Burford,N., Baughn,M.R., Azimzal,Y., Lu,D.A., Au-Young,J. and
Patterson,C.

TITLE Human transport proteins
JOURNAL Patent: WO 0078953-A 46 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1. 1826
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 415 a 519 c 512 g 380 t
ORIGIN

Query Match 99.9%; Score 1231; DB 6; Length 1826;
Best Local Similarity 100.0%; Pred. No. 7.6e-268;
Matches 1231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTTAGGCGCCAGGAGACGCCGAGCCTTACCTGTCGCCGGGACAGAGTCTTACCCA 61
Db 1 CCTTAGGCGCCAGGAGACGCCGAGCCTTACCTGTCGCCGGGACAGAGTCTTACCCA 60
QY 62 CCCAGTCTGTGTTCTGACGCCCCAGCTCATTCGCCAAATTAGGCTTGGGCTGGCTT 121
Db 61 CCCAGTCTGTGTTCTGACGCCCCAGCTCATTCGCCAAATTAGGCTTGGGCTGGCTT 120
QY 122 GTTCCCTCCGGCTCGAACCACTCTTCTCTGAGCCGAGCCAGCTACCGGGCTCTGGA 181
Db 121 GTTCCCTCCGGCTCGAACCACTCTTCTCTGAGCCGAGCCAGCTACCGGGCTCTGGA 180
QY 182 ATTGCCACCCCTCCCTGGGACCCCTTGAGGCTCCGTTGAGGAGCTACGGGGCAGAGC 241
Db 181 ATTGCCACCCCTCCCTGGGACCCCTTGAGGCTCCGTTGAGGAGCTACGGGGCAGAGC 240
QY 242 GGGACGTGAGCCTGAGTTGCTGACGGGCTCTCTGTGTGTGGTGGCTTCCCAATC 301
Db 241 GGGACGTGAGCCTGAGTTGCTGACGGGCTCTCTGTGTGTGGTGGCTTCCCAATC 300
QY 302 CCCGTGCCACCGGGGTGGGCGCGGGGAAAGCTCCTGCCCTCCCTGCTGGTCGGCTC 361
Db 301 CCCGTGCCACCGGGGTGGGCGCGGGGAAAGCTCCTGCCCTCCCTGCTGGTCGGCTC 360
QY 362 ACGGTGAGCTCCCGGTGATGGCTGGAGGGCCCCGGCGACAGCGGAGAGAGAGG 421
Db 361 ACGGTGAGCTCCCGGTGATGGCTGGAGGGCCCCGGCGACAGCGGAGAGAGAGG 420
QY 422 AAGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGA 481
Db 421 AAGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGA 480
QY 482 CATCAACATCCAGGAACCTCGCTGGGACAAAGTACTTCTCGGCGAGAGCCCGGCACTT 541
Db 481 CATCAACATCCAGGAACCTCGCTGGGACAAAGTACTTCTCGGCGAGAGCCCGGCACTT 540
QY 542 TTTCACTGTTACTGATCCTGAAATCTGCTGCTGCCGGGACAGCTGGAAGCTTCTCG 601
Db 541 TTTCACTGTTACTGATCCTGAAATCTGCTGCTGCCGGGACAGCTGGAAGCTTCTCG 600
QY 602 GAACATCGTGAGAACTACAGGGCGGGCTGTGACCCAGGAGTACCGAGAGCAGCT 661
Db 601 GAACATCGTGAGAACTACAGGGCGGGCTGTGACCCAGGAGTACCGAGAGCAGCT 660
QY 662 GTGAGGGGCAAGTATGTATGACTCCGCTTCCATCCGAGACAGGGGAGAAGTGGT 721
Db 661 GTGAGGGGCAAGTATGTATGACTCCGCTTCCATCCGAGACAGGGGAGAAGTGGT 720
QY 722 CCTGATGGCCGATGTCAGCCAGGTGCCCATGAACATGACCATCAGCTGCTGCTGCT 781
Db 721 CCTGATGGCCGATGTCAGCCAGGTGCCCATGAACATGACCATCAGCTGCTGCTGCT 780
QY 782 CACATTTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCTTCAA 841
Db 781 CACATTTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCTTCAA 840
QY 842 TGGCATTTAACTACTCCAAACGCGAGTGTGACACTCCCATCAGTGTGAGGCGAGCTGGG 901
Db 841 TGGCATTTAACTACTCCAAACGCGAGTGTGACACTCCCATCAGTGTGAGGCGAGCTGGG 900

Db 841 TGGCATTTAACTACTCCAAACGCGAGTGTGACACTCCCATCAGTGTGAGGCGAGCTGGG 900
QY 902 GACAGCCCTATGTAGTGGCCACCAGCTGAGCTGTGGCCAGCGCCCTGGGACTCAATCCCT 961
Db 901 GACAGCCCTATGTAGTGGCCACCAGCTGAGCTGTGGCCAGCGCCCTGGGACTCAATCCCT 960
QY 962 CACCAAGCACCTGCCCCCTTGGTCGCGAGATTGTGCCCCCTTGCACAGTGGCAGCTGC 1021
Db 961 CACCAAGCACCTGCCCCCTTGGTCGCGAGATTGTGCCCCCTTGCACAGTGGCAGCTGC 1020
QY 1022 CAACTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGAGTGGCATCCCGGTGGC 1081
Db 1021 CAACTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGAGTGGCATCCCGGTGGC 1080
QY 1082 TGATGAGGAGGTGAGAGGCTTGGCTACTCGGTGACTGACAGCCAAAGGAGAACTTCCA 1141
Db 1081 TGATGAGGAGGTGAGAGGCTTGGCTACTCGGTGACTGACAGCCAAAGGAGAACTTCCA 1140
QY 1142 GGTGATGATTCAAGAATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCAT 1201
Db 1141 GGTGATGATTCAAGAATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCAT 1200
QY 1202 GGACACTCTGAGAGAAGAAAGACTTCTGAAG 1232
Db 1201 GGACACTCTGAGAGAAGAAAGACTTCTGAAG 1231

RESULT 2
AX188304/c 3047 bp DNA linear PAT 06-AUG-2001
LOCUS AX188304
DEFINITION Sequence 3999 from Patent WO0142467.
ACCESSION AX188304
VERSION AX188304.1 GI:15139777
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
TITLE Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3999 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
source 1. 3047
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 629 a 790 c 891 g 733 t 4 others
ORIGIN

Query Match 98.1%; Score 1208.2; DB 6; Length 3047;
Best Local Similarity 99.8%; Pred. No. 1.1e-262;
Matches 1210; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 20 GCCGAGCGTTACCTGGTCCCGGGCAGCGGAGTCTTTTACCACCCCACTTGTGTTCTGA 79
Db 20 GCCGAGCGTTACCTGGTCCCGGGCAGCGGAGTCTTTTACCACCCCACTTGTGTTCTGA 79
QY 80 CGCCCTAGCTCATTCGCCAAATTAGGGCTGGGCTGCTGCTGTTCCCTCGGCTCGAA 139
Db 80 CGCCCTAGCTCATTCGCCAAATTAGGGCTGGGCTGCTGCTGTTCCCTCGGCTCGAA 139
QY 140 CCACCTTTCTGTGAGCCGAGCCAGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGG 199
Db 140 CCACCTTTCTGTGAGCCGAGCCAGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGG 199
QY 2915 CCACCTTTCTGTGAGCCGAGCCAGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGG 2856
Db 2915 CCACCTTTCTGTGAGCCGAGCCAGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGG 2856
QY 200 GCACCTTGAGGCTCCGTGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGAGTT 259
Db 200 GCACCTTGAGGCTCCGTGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGAGTT 259
QY 2855 GCACCTTGAGGCTCCGTGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGAGTT 2796
Db 2855 GCACCTTGAGGCTCCGTGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGAGTT 2796
QY 260 TGCTGACAGGCGTCTGTGTGTGTGGTGGTCTTCCCAATCCCGTGCACCGGGTGG 319
Db 260 TGCTGACAGGCGTCTGTGTGTGTGGTGGTCTTCCCAATCCCGTGCACCGGGTGG 319

|||||
Db 2795 TGCTGAGGCGTGTCTGTGTGGTGGTTCTGCCAATCCCGTGCCACCGGGTGG 2736
QY 320 GCGGCGCGGGAAGCTCTGCCCCCTGCTGCTGCGCGTCAAGCGTCCGCGT 379
Db 2735 GCGGCGCGGGAAGCTCTGCCCCCTGCTGCTGCGCGTCAAGCGTCCGCGT 2676
QY 380 GATGGCTGGAGGGCCCGCGCGAGCGGAGGAGAGAGAGAGAGAGCGGCTTGAGAGCT 439
Db 2675 GATGGCTGGAGGGCCCGCGCGAGCGGAGGAGAGAGAGAGAGCGGCTTGAGAGCT 2616
QY 440 TCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACC 499
Db 2615 TCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACC 2556
QY 500 TCGCTGGAGCCAAAGTACTTTCTTGGCGAGAGCCCGGACCTTTTCACTGTTACTGATCC 559
Db 2555 TCGCTGGAGCCAAAGTACTTTCTTGGCGAGAGCCCGGACCTTTTCACTGTTACTGATCC 2496
QY 560 TCGAAATCTGCTGCTGCTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTA 619
Db 2495 TCGAAATCTGCTGCTGCTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTA 2436
QY 620 CAGGCGCGCGCTGTGTGACCCCAAGGATCACCGAGACACAGCTGTGAGGGCCAGATATGT 679
Db 2435 CAGGCGCGCGCTGTGTGACCCCAAGGATCACCGAGACACAGCTGTGAGGGCCAGATATGT 2376
QY 680 GTATGACTCCGCTTCCATCCGAGACACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTC 739
Db 2375 GTATGACTCCGCTTCCATCCGAGACACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTC 2316
QY 740 AGCCAGAGTGCCCATGACATGACATCACTGGCTGCATGCTCACTTCTACAGGAAGAC 799
Db 2315 AGCCAGAGTGCCCATGACATGACATCACTGGCTGCATGCTCACTTCTACAGGAAGAC 2256
QY 800 CCCAACCGTGGTGTCTTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTTAATCTACTC 859
Db 2255 CCCAACCGTGGTGTCTTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTTAATCTACTC 2196
QY 860 CAACCGCAGTGTGACACTCCCATCACTGTGAGGAGCTGGGACAGCCCTATGTGAGTGC 919
Db 2195 CAACCGCAGTGTGACACTCCCATCACTGTGAGGAGCTGGGACAGCCCTATGTGAGTGC 2136
QY 920 CACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCAACCAAGCACCCTGCCCC 979
Db 2135 CACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCAACCAAGCACCCTGCCCC 2076
QY 980 CTTGGTGGCAGATTTGTGCCCCCTTGACAGAGTGAGCGAGCTGCCAATGCATCAACATCCC 1039
Db 2075 CTTGGTGGCAGATTTGTGCCCCCTTGACAGAGTGAGCGAGCTGCCAATGCATCAACATCCC 2016
QY 1040 CCTGATGAGCGAGAGAGAGAGCTGAGAGTGGGCATCCCGGTGGCTGATGAGGAGGTCAGAG 1099
Db 2015 CCTGATGAGCGAGAGAGAGAGCTGAGAGTGGGCATCCCGGTGGCTGATGAGGAGGTCAGAG 1956
QY 1100 GCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTTCAGGTGGTGAATTCAGGAAT 1159
Db 1955 GCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTTCAGGTGGTGAATTCAGGAAT 1896
QY 1160 CTGCATGGCGATTCTGCGCATGGCCATCCCAACCATGATCAATGACACTCTGAGAGAGAA 1219
Db 1895 CTGCATGGCGATTCTGCGCATGGCCATCCCAACCATGATCAATGACACTCTGAGAGAGAA 1836
QY 1220 AGACTTCTCTGAAG 1232
Db 1835 AGACTTCTCTGAAG 1823

RESULT 3
BD127742 2702 bp DNA linear PAT 18-SEP-2002
LOCUS BD127742
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127742

VERSION BD127742.1 GI:23222687
KEYWORDS JP 2002017375-A/3173.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2702)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 3173 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3173
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10', C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
location/Qualifiers
FT CDS (70)..(1044).
location/Qualifiers
source 1..2702
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 608 a 757 c 699 g 638 t
ORIGIN
Query Match 68.9%; Score 848.8; DB 6; Length 2702;
Best local Similarity 99.8%; Pred. No. 2.1e-181;
Matches 850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 381 ATGCTGGAGAGGGCCCGCGCGAGACGGAGGAGAGAGAGAGCGGGTCTGAGAGCTT 440
Db 1 ATGCTGGAGAGGGCCCGCGCGAGACGGAGGAGAGAGAGAGAGCGGGTCTGAGAGCTT 60
QY 441 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCT 500
Db 61 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCT 120
QY 501 CGCTGGGACCAAGTACTTTCTGTGGCGAGAGCCCGGACCTTTTCACTGTTACTGATCCT 560
Db 121 CGCTGGGACCAAGTACTTTCTGTGGCGAGAGCCCGGACCTTTTCACTGTTACTGATCCT 180
QY 561 CGAAATCTGCTGCTGTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 620
Db 181 CGAAATCTGCTGCTGTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 240
QY 621 AGGCGCGCGTGTGACCCCAAGGATCACCGAGACACAGCTGTGAGGGCCAGATATGTG 680
Db 241 AGGCGCGCGTGTGACCCCAAGGATCACCGAGACACAGCTGTGAGGGCCAGATATGTG 300
QY 681 TATGACTCCGCTTCCATCCGAGACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTCA 740
Db 301 TATGACTCCGCTTCCATCCGAGACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTCA 360
QY 741 GCCCAGGTGCCCATGACATGACATCACTGGCTGCATGCTCAATCTACAGGAAGACC 800
Db 361 GCCCAGGTGCCCATGACATGACATCACTGGCTGCATGCTCAATCTACAGGAAGACC 420
QY 801 CCAACCGTGGTGTCTGTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTGTAATCTACTCC 860
Db 421 CCAACCGTGGTGTCTGTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTGTAATCTACTCC 480

QY 861 AACCGAGTGTGACACTCCCATCTGTTGAGGCAGCTGGGGACACCTATGTGAGTCCC 920
|||||
Db 481 AACCGAGTGTGACACTCCCATCTGTTGAGGCAGCTGGGGACACCTATGTGAGTCCC 540
QY 921 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTACCAAGCACCTGCCCCC 980
|||||
Db 541 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTACCAAGCACCTGCCCCC 600
QY 981 TTGGTCGGCAGATTGTGCCCCCTTTGACAGCAGTGGCAGCTGCCCACTGCATCAACATCCCC 1040
|||||
Db 601 TTGGTCGGCAGATTGTGCCCCCTTTGACAGCAGTGGCAGCTGCCCACTGCATCAACATCCCC 660
QY 1041 CTGATGAGGCAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGGCAGGTCAAGAG 1100
|||||
Db 661 CTGATGAGGCAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGGCAGGTCAAGAG 720
QY 1101 CTGGCTACTCGGTGACTGCAGCCAGCAGGGAATCTTCCAGTGGTGAATTCAGAATC 1160
|||||
Db 721 CTGGCTACTCGGTGACTGCAGCCAGCAGGGAATCTTCCAGTGGTGAATTCAGAATC 780
QY 1161 TGCATGGCGATTCTCTGCCATGGCCATCCCACTGATCATGAGACTCTGGAGAGAAA 1220
|||||
Db 781 TGCATGGCGATTCTCTGCCATGGCCATCCCACTGATCATGAGACTCTGGAGAGAAA 840
QY 1221 GACTTCCTGAAG 1232
|||||
Db 841 GACTTCCTGAAG 852

RESULT 4
AK074707

LOCUS 2702 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90226 fis, clone NT2RM100858, weakly similar
to Tricarboxylate carrier.
AK074707

ACCESSION AK074707.1 GI:22760323
VERSION
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2702)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hrl.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..2702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM100858"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM1"
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2

FEATURES
source

CDS
neuronal precursor cells"
70..1047
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC1151.1"
/db_xref="GI:22760324"
/translation="MESKMGELPLDINIOEPRMDOSTPLGRHFFVTYDPRNLLSG
ADLEASRNIVQNYRAGVVTPEGTEDQLMRAKYVYDSAFHDTGEKVVILGRMSAOVPM
NMRTTGMLEFYRKPTVVFQWVNFQNSFNAIVNSRSGDPTITVROLGTAVASATG
AVATALGLKSLTKHLPPLVGRVPEFAVAANCINIPLMQRELOVGIPIVADEAGRL
GYSVTAAKQGIQVVISRICMAIPAMAPLIMDTLEKDELKRRPWLGAFLQVGLVG
FCLVFATPLCALFPQKSSIHISNLEPELRAQIHQNPSEVYVYNNKGL"
BASE COUNT 608 a 757 c 699 g 638 t
ORIGIN

Query Match 68.9%; Score 848.8; DB 9; Length 2702;
Best Local Similarity 99.8%; Pred. No. 2,1e-181;
Matches 850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 381 ATGGCTGGAGGCGCCGCGGCGAGCAGCGAGGAGAGAGAGGCGGCTTCTGAGAGCTT 440
|||||
Db 1 ATGGCTGGAGGCGCCGCGGCGAGCAGCGAGGAGAGAGAGGCGGCTTCTGAGAGCTT 60
QY 441 CAGAGAGCATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCT 500
|||||
Db 61 CAGAGAGCATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCT 120
QY 501 CGCTGGACCAAGTACTTCTCTGGGAGAGACCCCGCACCTTTTCACTGTACTGATCCT 560
|||||
Db 121 CGCTGGACCAAGTACTTCTCTGGGAGAGACCCCGCACCTTTTCACTGTACTGATCCT 180
QY 561 CGAATCTGCTGCTGTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 620
|||||
Db 181 CGAATCTGCTGCTGTCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 240
QY 621 AGGGCCGCGGTGTGACCCCGAGGATCACCGAGGACAGCAGCTGTGGAGGGCCAAATATGTG 680
|||||
Db 241 AGGGCCGCGGTGTGACCCCGAGGATCACCGAGGACAGCAGCTGTGGAGGGCCAAATATGTG 300
QY 681 TATGACTCCGCTTCCATCCGGACACAGGGGAGAGGTGCTGATTTGGCCGATGTCA 740
|||||
Db 301 TATGACTCCGCTTCCATCCGGACACAGGGGAGAGGTGCTGATTTGGCCGATGTCA 360
QY 741 GCCCAGTCCCATGAACATGACCATGCTGCTGATGCTCATCTTACAGGAAGACC 800
|||||
Db 361 GCCCAGTCCCATGAACATGACCATGCTGCTGATGCTCATCTTACAGGAAGACC 420
QY 801 CCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTACTACTCC 860
|||||
Db 421 CCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTACTACTCC 480
QY 861 AACCGAGTGTGACACTCCCATCACTGTGAGGCAGCTGGGGACACCTATGTGAGTCCC 920
|||||
Db 481 AACCGAGTGTGACACTCCCATCACTGTGAGGCAGCTGGGGACACCTATGTGAGTCCC 540
QY 921 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTACCAAGCACCTGCCCCC 980
|||||
Db 541 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTACCAAGCACCTGCCCCC 600
QY 981 TTGGTCGGCAGATTGTGCCCCCTTTGACAGCAGTGGCAGCTGCCCACTGCATCAACATCCCC 1040
|||||
Db 601 TTGGTCGGCAGATTGTGCCCCCTTTGACAGCAGTGGCAGCTGCCCACTGCATCAACATCCCC 660
QY 1041 CTGATGAGGCAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGGCAGGTCAAGAG 1100
|||||
Db 661 CTGATGAGGCAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGGCAGGTCAAGAG 720
QY 1101 CTGGCTACTCGGTGACTGCAGCCAGCAGGGAATCTTCCAGTGGTGAATTCAGAATC 1160
|||||
Db 721 CTGGCTACTCGGTGACTGCAGCCAGCAGGGAATCTTCCAGTGGTGAATTCAGAATC 780
QY 1161 TGCATGGCGATTCTCTGCCATGGCCATCCCACTGATCATGAGACTCTGGAGAGAAA 1220
|||||

Db	781	TGCATGGCGATTCCCTGCGCATGGCCATCCACCACTGATCATGGACACTCTGGAGAGAAA	840
QY	1221	GACTTCCTGAAG 1232	
Db	841	GACTTCCTGAAG 852	
RESULT 5			
LOCUS	BC000124	2735 bp mRNA linear PRI 12-JUL-2001	
DEFINITION		Homo sapiens, similar to tricarboxylate carrier-like protein, clone	
ACCESSION	BC000124	IMAGE:3352015, mRNA, complete cds.	
VERSION	BC000124.1	GI:12652746	
KEYWORDS	MG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	1 (bases 1 to 2735)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK		Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
COMMENT		NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabh, Parvaneh Saeedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zyderdun, Marco Marra.	
FEATURES		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAL Plate: 6 Row: k Column: 13 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction. Location/Qualifiers 1. 2735 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:3122 IMAGE:3352015" /tissue_type="Eye, retinoblastoma" /clone_id="NIH-MGC_16" /lab_host="DH10B-R" /note="Vector: POTB7" 75. 1040 /codon_start=1 /product="Similar to tricarboxylate carrier-like protein" /protein_id="AAH00124.1" /db_xref="GI:12652747" /translation="MGELPLDINIOEPRWDSTFLGARHFFVTIDPRNLLSGAOLEASRNIVQNRAGVVTGITEQDLNRAKYVDSAFHPDTGEKVVILGRMSAQVPMNMTITGCMLETRYRKTPTVFEWQWVNOSEFNIVNSRSGDPTITVRQLGTRAYVSATGAVATALGLSLTKHLPLPLVGRFVFPFAVAANCINIPLMRQRELQVIGIPVADAGQRLGYSVTAAGQIFQVVISRICMAIPAMAIPPLIMDTLEKKDFLKRPRWGLAPLQVGLVGFCLVFAIPALCALFPQKSSIHISNLEPELRAQIHEONPSVEVYYNKGIL"	
CDS			
BASE COUNT	645	a 757	c 695 g 638 t
ORIGIN			

Query Match	Best Local	Similarity	Score	DB	Length
Matches	845;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
68.68;	100.0%;	Pred. No. 1.5e-180;			
388	GGAGGGCCCCGGCGGACAGCGGAGGCGAGAGGAAAGCGGCTTGAGAGCTTCAGAGAG	447			
1	GGAGGGCCCCGGCGGACAGCGGAGGCGAGAGGAAAGCGGCTTGAGAGCTTCAGAGAG	50			
448	CGATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAACTCGCTGGG	507			
61	CGATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAACTCGCTGGG	120			
508	ACCAAAAGTACTTTCCTGGGACAGCCCCGGCAGCTTTTCACTGTGTACTGATCCTCGAATC	567			
121	ACCAAAAGTACTTTCCTGGGACAGCCCCGGCAGCTTTTCACTGTGTACTGATCCTCGAATC	180			
568	TGCTGCTGTCCGGGACACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCG	627			
181	TGCTGCTGTCCGGGACACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCG	240			
628	GGGTGTGACCCCGAGGATCACCGAGGAGCCAGCTGTGGAGGCGCAAGTATGTATGACT	687			
241	GGGTGTGACCCCGAGGATCACCGAGGAGCCAGCTGTGGAGGCGCAAGTATGTATGACT	300			
688	CCGCTTCCATCCGGACACAGGGGAGAAAGTGTCTGATTTGGCCGATGTACAGCCAGG	747			
301	CCGCTTCCATCCGGACACAGGGGAGAAAGTGTCTGATTTGGCCGATGTACAGCCAGG	360			
748	TGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACG	807			
361	TGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACG	420			
808	TGCTGTCTTCTGGCAGTGGGTGAATCAGTCTCTCAATGCCATTTTAACTACTCCAAACGCA	867			
421	TGCTGTCTTCTGGCAGTGGGTGAATCAGTCTCTCAATGCCATTTTAACTACTCCAAACGCA	480			
868	GTGTGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCTTATGTGAGTGCCACCACTG	927			
481	GTGTGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCTTATGTGAGTGCCACCACTG	540			
928	GAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACTGCCCTTGTGCTG	987			
541	GAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACTGCCCTTGTGCTG	600			
988	GCAGATTTGTGCCCTTTGACAGCAGTGCGAGCTGCCAACTGCATCAACATCCCTGATGA	1047			
601	GCAGATTTGTGCCCTTTGACAGCAGTGCGAGCTGCCAACTGCATCAACATCCCTGATGA	660			
1048	GGCAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGCGAGTCCAGAGGCTTGGCT	1107			
661	GGCAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGCGAGTCCAGAGGCTTGGCT	720			
1108	ACTCGGTGACTGCAGGCCAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGAATCTGCATGG	1167			
721	ACTCGGTGACTGCAGGCCAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGAATCTGCATGG	780			
1168	CGATTCTTGCCATGGCCATCCCACTGATCATGAGACATCTGGAGAAAGAACTTCC	1227			
781	CGATTCTTGCCATGGCCATCCCACTGATCATGAGACATCTGGAGAAAGAACTTCC	840			
1228	TGAAG 1232				
841	TGAAG 845				

SOURCE	ORGANISM	Homo sapiens (human)
REFERENCE	AUTHORS	Mintz, L., Freilich, S. and Bernstein, J.
TITLE	Novel nucleic acid and amino acid sequences	
JOURNAL	Patent: WO 0206315-A 115 24-JAN-2002;	
FEATURES	CompuGen Ltd. (IL)	
source	location/Qualifiers	
	1..984	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
BASE COUNT	224 a 278 c 285 g 196 t 1 others	
ORIGIN		
Query Match	Best Local Similarity	66.4%; Score 818.6; DB 6; Length 984;
Matches	Conservative	829; 1; Mismatches 0; Indels 1; Gaps 1;
QY	402 CGACAGCGGAGCAGAGAGAAGCGGTCTGAGAGCTTCAGAGCGATGGAAGCAA	461
Db	1 CGACAGCGGAGCAGAGAGAAGCGGTCTGAGAGCKTCAAGAGCGATGGAAGCAA	60
QY	462 ATGGTGAAATTGCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCACAAGTACTTTC	521
Db	61 ATGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCACAAGTACTTTC	120
QY	522 CTGGGACAGAGCCCGGCACTTTTCACTGTTACTGATCCTCGAATCTGCTGCTGCCGG	581
Db	121 CTGGGACAGAG - CCGGCACTTTTCACTGTTACTGATCCTCGAATCTGCTGCTGCCGG	179
QY	582 GCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGGCTGTGACCCCA	641
Db	180 GCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGGCTGTGACCCCA	239
QY	642 GGGATCACCGAGGACCACTGTGGAGGGCCCAAGTAGTAGTACTCCGCCCTTCCATCCG	701
Db	240 GGGATCACCGAGGACCACTGTGGAGGGCCCAAGTAGTAGTACTCCGCCCTTCCATCCG	299
QY	702 GACACAGGGGAGAAAGTGTCTGATTGGCCCATGTCCAGCCCAAGTGGCCCATGAACATG	761
Db	300 GACACAGGGGAGAAAGTGTCTGATTGGCCCATGTCCAGCCCAAGTGGCCCATGAACATG	359
QY	762 ACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTCTCGCAG	821
Db	360 ACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTCTCGCAG	419
QY	822 TGGGTGAATCAGTCTCTTCATGCCATTTGTTAACTACTCCAACCCGCAAGTGTGACACTCCC	881
Db	420 TGGGTGAATCAGTCTCTTCATGCCATTTGTTAACTACTCCAACCCGCAAGTGTGACACTCCC	479
QY	882 ATCACTGTGAGGACAGCTGGGGGACAGCCATATGTAGTGGCACCACCTGGAGCTGTGGCCACG	941
Db	480 ATCACTGTGAGGACAGCTGGGGGACAGCCATATGTAGTGGCACCACCTGGAGCTGTGGCCACG	539
QY	942 GCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCCCTTGGTTCGGCAGATTTGTGCC	1001
Db	540 GCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCCCTTGGTTCGGCAGATTTGTGCC	599
QY	1002 TTTCAGCAGTGGCAGCTGSCCAACTGCATCAACATCCCTGATGAGGCAGAGAGAGAGCTG	1061
Db	600 TTTCAGCAGTGGCAGCTGSCCAACTGCATCAACATCCCTGATGAGGCAGAGAGAGAGCTG	659
QY	1062 CAGGTGGGATCCCGTGGCTGATGAGGCAGGTCAAGAGGCTTGGCTACTCGGTGACTGCA	1121
Db	660 CAGGTGGGATCCCGTGGCTGATGAGGCAGGTCAAGAGGCTTGGCTACTCGGTGACTGCA	719
QY	1122 GCCAAGCAGGGAATCTTCCAGGTGTGATTTCAAGAATCTGCATGGCGATTCTTGCCATG	1181
Db	720 GCCAAGCAGGGAATCTTCCAGGTGTGATTTCAAGAATCTGCATGGCGATTCTTGCCATG	779

QY	1182	GCCATCCCACCACCTGATCATGCAGACTCTGGAGAAGAAGACTTCCTGAAG	1232
Bc012208			
LOCUS			
DEFINITION	Bc012208	2870 bp	mRNA linear ROD 16-APR-2003
ACCESSION	Mus musculus sideroflexin 3, mRNA (cdna clone MGC:18958		
VERSION	IMAGE:3985155), complete cds.		
KEYWORDS	BC012208		
SOURCE	BC012208.1	GI:15126553	
ORGANISM	MGC.		
	Mus musculus (house mouse)		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2870)		
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenman,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahney,J., Helton,E., Kettaman,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 12477932 2 (bases 1 to 2870) Strausberg,R. Direct Submission Submitted (06-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov/ Contact: MGC help desk Email: cgapsb-email.nih.gov Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgebcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
REMARK			
COMMENT			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source			
	1. .2870		
	/organism="Mus musculus"		
	Location/Qualifiers		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.linl.gov Series: IRAK Plate: 23 Row: 1 Column: 8 This clone was selected for full length sequencing because it passed the following selection criteria: GenomScan gene prediction, Similarity but not identity to protein.		


```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:18958 IMAGE:3985155"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_1lb="NCI_CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 2870
/gene="Sfxn3"
/db_xref="LocusID:94280"
/db_xref="MGI:2137679"
221. 1186
/codon_start=1
/product="sideroflexin 3"
/protein_id="AAH12208.1"
/db_xref="GI:15126554"
/db_xref="LocusID:94280"
/translacion="MGDPLPINIQEPWDQSTFLGRARHFTVTPRNNLLSGEQLE
ASRNIVONTYRAGVATPGLTEDQNLWRAKYVDSAFHPDITGEKVVLLIGMSAQVPMNMTI
TGCMLTFYRKPTVVFQWQVNOSENAIVNSNRSDAPITVQQLGTAVSATTGAVAT
ALGLKSLTKHLPPLVGRFVPEAFAVAANCINIPLMRQRELQVGIPTVDEAGQRLGHSV
TAAKGQIGQVVISRIGMAIPMAIPVYIMNTLEKKDFLKRPRWLGAPLOYGLVGFCV
FATPLCALFPQRSSIHVTRLEPELRAQIOAQNPSIDVYVYNNKGL"
BASE COUNT      622 a      804 c      696 g      748 t
ORIGIN

```

Query Match	54.78;	Score 674;	DB 10;	Length 2870;
Best Local Similarity	85.48;	Pred. No. 7.2e-142;		
Matches 786;	Conservative 0;	Mismatches 130;	Indels 4;	Gaps 3;
QY 313	CGGTTGGCGCGCGCGGAAAGCTCCTGCCCTCCCTGCTGTGCGCGTACGCGTGACGT	372		
Db 76	CAGCGAACGAGCTCAGGAGCGCCCGCCCTCCCTGCTGCTCAGCGTACGCGTGACGT	135		
QY 373	CCCGCGTATGCGCTGGGAGGGCGCGCGGACAGCGGAGGACAGAGGAGCGCGTTCT	432		
Db 136	CTCG-GTGATGCGCTGGGAGGAAA--GCCGAGAGCGGTGAGGAGCGCGGTCTGAGAG-CT	191		
QY 433	GAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAATCC	492		
Db 192	TCTAGAGGCTGAAACCCCGGAAAGCAGATGGGTGACCTGCCCTTAATATCAACATCC	251		
QY 493	AGGAACCTCGCTGGGACCAAGTACTTTCCTGGGCAAGCGCCGACATTTCACGTGTA	552		
Db 252	AGGAACCTCGGTGGGACCAAGCACATTTCAGGACAGAGCCCGCATTTCTTCACAGTCA	311		
QY 553	CTGATCCTCGAAATCTGCTGCTGTCCGGGGACAGCTGGAAGCTTCTCGAACATCGTGC	612		
Db 312	CTGATCCCGGAAATCTGCTGCTGTCCGGGGACAGCTGGAAGCTTCCCGAACATCGTGC	371		
QY 613	AGAACTACAGGCGCGCGGTGGTGACCCAGGATCACCGAGACAGCTGNGAGGGCCA	672		
Db 372	AGAAATACAGGCGTGTGTGGCAACCCCGGTCTCACTGAGGACCAAGCTATGGCGAGCCA	431		
QY 673	AGTATGTGATGACTCCGCGCTTCATCCGGACACAGGGGAGAAGGTGCTCATTTGGCC	732		
Db 432	AATACGTGATGACTCAGCATTCATCCGGACACAGGGGAGAAGGTGCTCATTTGGCC	491		
QY 733	GCATGTACGCCCAAGGTGCCCATGAAACATGACATCFACTGGCTGCATGCTACATTCTACA	792		
Db 492	GTATGTACGCCCAAGGTGCCCATGAAACATGACATTAATTGGCTGCATGCTACCTTCTACA	551		
QY 793	GGAAGACCCCAACCGTGTGTCTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTA	852		
Db 552	GGAAGACTCCGACTGTGTGTCTCTGGCAGTGGGTCAATCAGTCCCTTCAATGCTATTGTGA	611		
QY 853	ACTACTCCAACCGCAGTGGTGACACTCCCATCACTGTGAGGCAGCTGGGGACAGCCTATG	912		
Db 612	ATTACTCTAATCGCAGCGCGGATGCTCCCATCACTGTGACAGCAGTTGGGACAGCCTATG	671		
QY 913	TGAGTGCCACACTGGAGCTGTGGCCACGGCGCTGGGACTCAAAATCCCTCAACCAAGCAC	972		

Db	672	TGAGTGCACCACTGGGGGTGTGGCTACTGCTCTGGGACTCAAGTCTCTCACCAAGCACC	731
QY	973	TGCCCCCTTGTCGGCGAGATTGTGCCCTTTGCAGCAGTGCCAGCTGCCAATGCATCA	1032
Db	732	TGCCCCCGCTAGTCGGTCGATTCGTGCCCTTTGCAGCTGTGGCCGCTGCCAACTGCATCA	791
QY	1033	ACATCCCCCTGATGAGCAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAG	1092
Db	792	ACATCCCCCTGATGAGCAGAGAGAGCTGCAGGTGGGCATCCAGTAGCATGAGGCTG	851
QY	1093	GTCAGAGGCTTGGCTACTCCGGTGACTGCAGCCCAAGCAGGAATCTTCCAGGTGTGATTT	1152
Db	852	GTCAGAGGCTTGGCCACTTCGGTGACTGCTGCCAACAGGGAATCTTCCAGGTGTGATAT	911
QY	1153	CAAGAATCTGCATGGCGCATTCCTGCCATGGCCATCCCACTGATCATGAGACTCTGG	1212
Db	912	CAAGAATCGGAATGGCGATTCGCCCATGGCCATTCGCCGCTGATCATGAACACTCTGG	971
QY	1213	AGAAGAAAGACTTCTCTGAAG	1232
Db	972	AGAAGAAAGACTTCTCTGAAG	991
RESULT	8		
AF276997		2699 bp	mRNA linear ROD 22-JUN-2000
LOCUS			
DEFINITION	AF276997	Rattus norvegicus tricarboxylate carrier-like protein mRNA,	
complete cds.			
ACCESSION	AF276997		
VERSION	AF276997.1	GI:8650525	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 (bases 1 to 2699)		
TITLE	Mashima,H. and Kojima,I.		
JOURNAL	Submitted (12-JUN-2000) Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15, Showa-machi, Maebashi, Gunma 371-8512, Japan		
FEATURES			
source			
CDS			
polyA_signal			
BASE COUNT			
ORIGIN			
Query Match	53.9%;	Score 664.2;	DB 10; Length 2699;
Best Local Similarity	82.5%;	Pred. No. 1.2e-139;	
Matches 785;	Conservative 0;	Mismatches 163;	Indels 3; Gaps 2;
Db	282	GTGGCTGGGTTCTGCCAATCCCCGTGCCACCGGGGTGGGGCGGGCCGGAAGCTCCTGCC	341
	56	GTCCCGGAAGCTTGCACAAGAGGGGGGTGGGTACAGGCGAACCAGTTTAGGGAAGCCCCGCC	115
QY	342	CCTCCCTGCTGCTGGCGGTCAACGCGGTACGCTCCC CGGTGATGGCTGGGAGGGCCCCGGCGG	401

Db 116 CTCCCTGCTGCTCAGCGCTCAGCGGTGACGCTCCGG-GTGATGGCTGGAGAAA--GCCG 172

QY 402 CGACAGCGGAGGCGAGAGAGAAAGCGGTTCTGAGAGCTTCAGAGAGCGATGAAAGCAAA 461

Db 173 AGAGCTGTAGAGAAAGCGGGTCTGAGAGCTGTAGAGGCTGTAAATCCCGGAAAGCAAG 232

QY 462 ATGGGTGAATGCTTTAGACATCAACATCCAGAACCTCGCTGGAGCAAGTACTTTC 521

Db 233 ATGGGCACTTGCCCTTAATATCAACATCCAGAACCTCGATGGAGCAAGACATTC 292

QY 522 CTGGGAGAGCCCGCACTTTTTCACCTGTACTGATCCTCGAAATCTGCTGCTGCCGG 581

Db 293 CTAGGCAAGAGCCCGCACTTCTTACAGTACTTCCCGAAATCTCCTGCTGCCGG 352

QY 582 GCACAGCTGAAGCTTCTCGGAACATCGTCGACACTACAGGGCCGCGGTGAGACCCA 641

Db 353 AAACAGTTGGAAGCTTCTCGGAACATCGTCGACAAATTACAGGGCTGGTGTAACTCCT 412

QY 642 GGGATCACCAGAGACAGCTGTGAGGGCCAGTATGTATGACTCCGCTTCATCCG 701

Db 413 GGGCTCACTGAGGACCACTATGAGAGGCCAAATACGTATGACTCAGACTTCCACCCG 472

QY 702 GACACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTCAGCCAGGTGCCATGAACATG 761

Db 473 GACACAGGGGAGAGAGTGGTCTGATTTGGCCGATGTCAGCCAGGTGCCATGAACATG 532

QY 762 ACCATCACTGGCTGCATGCTCACAATTTACAGAGAACCCCAACCGTGGTCTTGGCAG 821

Db 533 ACCATTACTGGCTGCATGCTCACAATTTACAGAGAACCTCCAACTGTGGTCTTGGCAG 592

QY 822 TGGGTGAATCAAGTCTTCAATGCCATTTGTAACACTTCCAAACCGAGTGGTGAACACTCC 881

Db 593 TGGGTGAATCAAGTCTTCAATGCCATTTGTAACACTTCCAAACCGAGTGGTGAACACTCC 652

QY 882 ATCACTGTGAGGAGAGCTGGGAGACAGCCATGTAGTGCACCACTGGAGCTGTGCCACG 941

Db 653 ATCACTGTGAGGAGAGCTGGGAGACAGCCATGTAGTGCACCACTGGAGCTGTGCCACT 712

QY 942 GCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGGTCGGCAGATTGTGCC 1001

Db 713 GCTCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCACTAGTGGGTGCAATTCGTACCC 772

QY 1002 TTGCAAGAGTGGCAGCTGCCAAGCTGCATCAACATCCCCCTGATGAGGAGAGAGCTG 1061

Db 773 TTGCAAGCTGTGGCTGCCCAAGCTGCATCAACATCCCCCTGATGAGGAGAGAGCTG 832

QY 1062 CAGGTGGGACTCCCGGTGGCTGATGAGGAGAGCTGAGGCTTGGCTACTCGTACTGCA 1121

Db 833 CAGGTGGGACTCCCGGTGGCTGATGAGGAGAGCTGAGGCTTGGCTACTCGTACTGCT 892

QY 1122 GCCAAGCAGGGAATCTCCAGGTGTGATTTCAAGAAATCTGCATGGCGATTCCTGCCATG 1181

Db 893 GCCAAGCAGGGAATCTCCAGGTGTGATTTCAAGAAATCGCATGGCGATTCCTGCCATG 952

QY 1182 GCCATCCCACTGATGATGAGACACTCTGAGAAAGAAAGACTTCTGAAG 1232

Db 953 GCCATTCCTCCCGGTGATGATGAGACACTCTGAGAAAGAAAGACTTCTGAAG 1003

RESULT 9

BD124792

LOCUS BD124792 752 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD124792

VERSION BD124792.1 GI:23219737

KEYWORDS JP 2002017375-A/223.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 752)

AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,

TITLE Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.

JOURNAL Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002017375-A 223 22-JAN-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002017375-A/223

PD 22-JAN-2002

PF 07-JUL-2000 JP 200253172

PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO

PI ISHII,

PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI

SHINICHI KOJIMA,

PI TETSUJI OTSUKI,HISASHI KOGA

PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC

10, C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

PC C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH key

location/Qualifiers

FT source 1.752

FT location/Qualifiers

1.752

/organism="Homo sapiens (human)".

location/Qualifiers

1.752

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 173 a 208 c 214 g 149 t 8 others

ORIGIN

Query Match 53.5%; Score 659; DB 6; Length 752;

Best Local Similarity 96.5%; Pred. No. 1.9e-138;

Matches 713; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 381 ATGGCTGGAGGGCCCGCGCGGACAGCGAGGAGAGAGAAAGCGGTTCTGAGAGCTT 440

Db 1 ATGGCTGGAGGGCCCGCGCGGACAGCGAGGAGAGAGAAAGCGGTTCTGAGAGCTT 60

QY 441 CAGAGAGCGATGAAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACT 500

Db 61 CAGAGAGCGATGAAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACT 120

QY 501 CGCTGGAGCAAGTACTTCTCTGGGAGAGAGCCCGGCACTTTTCACTGTACTGATCCT 560

Db 121 CGCTGGAGCAAGTACTTCTCTGGGAGAGAGCCCGGCACTTTTCACTGTACTGATCCT 180

QY 561 CGAATCTGCTGCTGCTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTAC 620

Db 181 CGAATCTGCTGCTGCTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTAC 240

QY 621 AGGGCCGGCTGTGACCCCA-GGGATCACCGAGGAGACAGCTGTGAGGGCCAAATATGT 679

Db 241 AGGGCCGGCTGTGACCCCAAGGATCACCGAGGAGACAGCTGTGAGGGCCAAATATGT 300

QY 680 GTATGACTCCGCTTCCATCCGGACACAGGGGAGAAAGTGGTCTGATTGGCCGATGTC 739

Db 301 GTATGACTCCGCTTCCATCCGGACACAGGGGAGAAAGTGGTCTGATTGGCCGATGTC 360

QY 740 AGCCAGGTGCCATGAACATGACCATGAGCTGCTGATGCTCACAATCTACAGGAAGAC 799

Db 361 AGCCAGGTGCCATGAACATGACCATGAGCTGCTGATGCTCACAATCTACAGGAAGAC 420

QY 800 CCCAAGCGTGTCTTCTGGCAGTGGGTGAATCAGTCCCTCAATGCAATGTAACTACTC 859

Db 421 CCCAAGCGTGTCTTCTGGCAGTGGGTGAATCAGTCCCTCAATGCAATGTAACTACTC 480

QY 860 CAACCGCAGTGTGACACTCCCATCACTGTGAGGAGAGCTGGGAGACAGCTATGTAGTGC 919

Db 481 CAACCGCAGTGTGACACTCCCATCACTGTGAGGAGAGCTGGGAGACAGCTATGTAGTGC 540

QY 920 CACCACGTGAGCTGTGGCCACGGCCCTGGGAGCTCAAAATCCCTCACCAAGCACTGCCCC 979

Db 541 CACCACGTGAGCTGTGGCCACGGNCTGGGAGCTCAAAATCCCTCACCAAGCACTGCCCC 600

QY 980 CTTCGTCGGCAGATTGTGCCCCCTTTGCAGCAGTGGCAGCTGCCAAC-TGCATCAACATCC 1038
DB 601 CTTCGTCNGCAGATTGTGCCCCCTTTGCAGCAGTGGCNGCTGCCCAACTTGATCAACATCC 660
QY 1039 CCTGATGAGGCAGAGAGAGCTGC-AGGTGGCATCCCGGTG--CTGATGAGCAGAGTC 1095
DB 661 CCTGATGAAGCANAANAACATGCAAGGTNGGCATCCCAATGGCTGATGAAGCAGAGTC 720
QY 1096 AGAGCTTGCTACTCGGT 1114
DB 721 AAAAGCTTGCTACTCCGT 739

RESULT 10
BD126228
LOCUS BD126228 752 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD126228
VERSION BD126228.1 GI:23221173
KEYWORDS JP 2002017375-A/1659.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 752)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 1659 22-JAN-2002;
HELIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/1659
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT source 1..752
FT location/Qualifiers
1..752
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 173 a 208 c 214 g 149 t 8 others
ORIGIN

Query Match 53.5%; Score 659; DB 6; Length 752;
Best Local Similarity 96.5%; Pred. No. 1.9e-138;
Matches 713; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 381 ATGGCTGGAGGGCCCCGGCGGACAGCGGAGGAGAGGAGGCGGTTCTGAGAGCTT 440
DB 1 ATGGCTGGAGGGCCCCGGCGGACAGCGGAGGAGAGGAGGCGGTTCTGAGAGCTT 60
QY 441 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACTT 500
DB 61 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACTT 120
QY 501 CGCTGGAGCAAAAGTACTTTCTGGGCGAGAGCCCGGACATTTTCACTGTACTGATCCT 560
DB 121 CGCTGGAGCAAAAGTACTTTCTGGGCGAGAGCCCGGACATTTTCACTGTACTGATCCT 180

QY 561 CGAATCTGCTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 620
DB 181 CGAATCTGCTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 240
QY 621 AGGGCCGGCGTGGTGACCCCA-GGGATCACCGAGGAGCAGCTGTGGAGGGCCAAATATGT 679
DB 241 AGGGCCGGCGTGGTGACCCCAAGGGGATCACCGAGGAGCAGCTGTGGAGGGCCAAATATGT 300
QY 680 GTATGATCCGCCCTTCCATCCGACAGAGGGAGAGGTGTCCTGATTTGGCCGATGTC 739
DB 301 GTATGATCCGCCCTTCCATCCGACAGAGGGAGAGGTGTCCTGATTTGGCCGATGTC 360
QY 740 AGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGAC 799
DB 361 AGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGAC 420
QY 800 CCCAACCGTGGTGTCTTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTGTTACTACTC 859
DB 421 CCCAACCGTGGTGTCTTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTGTTACTACTC 480
QY 860 CAACCGAGTGTGTACACACTCCCATCACTGTGAGGACAGCTGGGAGACCCATATGTGAGTGC 919
DB 481 CAACCGAGTGTGTACACACTCCCATCACTGTGAGGACAGCTGGGAGACCCATATGTGAGTGC 540
QY 920 CACCAGTGAAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCC 979
DB 541 CACCAGTGAAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCC 600
QY 980 CTTCGTCGGCAGATTGTGCCCCCTTTGCAGCAGTGGCAGCTGCCAAC-TGCATCAACATCC 1038
DB 601 CTTCGTCNGCAGATTGTGCCCCCTTTGCAGCAGTGGCNGCTGCCCAACTTGATCAACATCC 660
QY 1039 CCTGATGAGGCAGAGAGAGCTGC-AGGTGGCATCCCGGTG--CTGATGAGCAGAGTC 1095
DB 661 CCTGATGAAGCANAANAACATGCAAGGTNGGCATCCCAATGGCTGATGAAGCAGAGTC 720
QY 1096 AGAGCTTGCTACTCGGT 1114
DB 721 AAAAGCTTGCTACTCCGT 739

RESULT 11
AF325262 2710 bp mRNA linear ROD 30-APR-2001
LOCUS AF325262
DEFINITION Mus musculus sideroflexin 3 (Sfxn3) mRNA, complete cds.
ACCESSION AF325262
VERSION AF325262.1 GI:13785615
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2710)
Fleming,M.D., Campagna,D.R., Haslett,J.N., Trenor,C.C. III and
Andrews,N.C.
TITLE A mutation in a mitochondrial transmembrane protein is responsible
for the pleiotropic hematological and skeletal phenotype of
flexed-tail (f/f) mice
JOURNAL Genes Dev. 15 (6), 652-657 (2001)
MEDLINE 21172735
PUBMED 11274051

REFERENCE 2 (bases 1 to 2710)
Fleming,M.D., Campagna,D.R. and Andrews,N.C.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Pathology, Children's Hospital, 300
Longwood Avenue, Boston, MA 02115, USA
Location/Qualifiers
1..2710

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2710
/gene="Sfxn3"


```

CDs
83. .1048
/gene="Sfxn3"
/note="SFXN3; similar to the flexed-tail (f) protein
sideroflexin 1 (SFXN1)"
/codon_start=1
/product="sideroflexin 3"
/protein_id="AAK39430.1"
/db_xref="GI:13785616"
/translation="MGDLPLNINIQEPKWDSTFLGRARHFTVTPRNLLSGEQL
ASRNIVQNYRAGVATPGLEEDQLAKYVYDSAFHPDGEKVLIGRMSAOVPMNTI
TSCMLTFYRKPTPTVFMQVNOSENAIVNYSNRSGDAPITVQQLGTAIVSATGAVAT
ALGLSLTKHLPLVGRFVPAVAANCINIPIMRQRELVGIPVTDGAGRLGHSY
TAAKQIFQVVISRIGMAIPMAIPVIMNTLEKKDLKRRPWLGAPLQVGLVGFCLV
FATPLCCALFPORSSIHVTRLEPRLAQIOAQNPSIDVYYNKGK"

BASE COUNT      591 a      755 c      647 g      711 t      6 others
ORIGIN

Query Match      53.2%; Score 655.6; DB 10; Length 2710;
Best Local Similarity 87.7%; Pred. No. 1e-137;
Matches 726; Conservative 3; Mismatches 92; Indels 7; Gaps 1;

QY 412 GGCAGAGAGGAAGCGGCTTGTAGAGCTT-----CAGAGAGCGATGGAAGCAAAATG 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 RCGGCTGAGGAAGCGGCTGTAGAGCTTCTAGAGCTGAAACCCCGGAAGCAAGATG 85

QY 465 GGTGAATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGGACCAAGTACTTTCCTG 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 86 GGTGACCTGCCCTTAATATCAACATCCAGGAACCTCGGTGGGACCAAGCATTTCTA 145

QY 525 GGCAGAGCGCGGCACTTTTCACTGTTACTGATCTCGAATCTGCTGCTCCGGGCA 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 146 GGCAGAGCGCGGCACTTTTCACTGTTACTGATCTCGAATCTGCTGCTCCGGGAA 205

QY 585 CAGCTGGAAGCTTCTCGGAACATCGTGCAAGAACTACAGGGCGCGTGTGACCCAGG 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 206 CAGCTGGAAGCTTCTCGGAACATCGTGCAAGAACTACAGGGCGCGTGTGACCAACCCGGGT 265

QY 645 ATCACCAGGACAGCTGTGAGGCGCAAGTATGTATGACTCCGCCCTTCATCCGAG 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 266 CTCACTGAGACAGCACTATGGCGCAAAATACGTATGATCACTACATCCGAG 325

QY 705 ACAGGGGAGAAGTGTGCTGATGGCCGATGTACGCCAGGTGCCATGAACATGACC 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 326 ACAGGGGAGAAGTGTGCTGATGGCCGATGTACGCCAGGTGCCATGAACATGACC 385

QY 765 ATCACTGGCTGCATGCTACATTTCTACAGGAAGACCCCAACCGTGTGCTGCGAGTGG 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 386 ATCACTGGCTGCATGCTACATTTCTACAGGAAGACTCCGACTGTGCTGCGAGTGG 445

QY 825 GTGAATCAGTCTTCAATGCCATTGTTAACTACTCCAACCGCAGTGTGACACTCCCATC 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 446 GTCAATCAGTCTTCAATGCCATTGTTAACTACTCCAATCGACGCGGATGCTCCCATC 505

QY 885 ACTGTGAGGACAGCTGGGAGACAGCTATGTAGTGGCCACCACTGAGCTGTGGCCAGGCC 944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 506 ACTGTGAGCAGTGTGGGAGACAGCTATGTAGTGGCCACCACTGAGCTGTGGCTACTGCT 565

QY 945 CTGGGACTCAATCCCTCACCAAGCACTGCCGCCCTTGGTGGGACAGATTGTGCCCTTT 1004
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 566 CTGGGACTCAATCCCTCACCAAGCACTGCCGCCCTTGGTGGGACAGATTGTGCCCTTT 625

QY 1005 GCAGCAGTGGCAGCTGCCAATGCATCAACATCCGCCCTGATGAGGACAGAGAGAGCTGCAG 1064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 626 GCAGCTGTGGCGCTGCCAATGCATCAACATCCGCCCTGATGAGGACAGAGAGAGCTGCAG 685

QY 1065 GTGGGACATCCCGGTGCTGATGAGGACAGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCC 1124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 686 GTGGGACATCCCGGTGCTGATGAGGACAGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCC 745

QY 1125 AAGCAGGGAATCTTCCAGGTGATGATTCAAGAATCTGCATGGCGATTCTCGCCATGGCC 1184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 746 AAGCAGGGAATCTTCCAGGTGATGATTCAAGAATCGGAATGGCGATTCTCGCCATGGCC 805

```

```

QY 1185 ATCCACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCCGTGAG 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 806 ATCCCGCGGTGATCATGACACTCTGAGAGAAGAAAGACTTCCGTGAG 853

```

```

RESULT 12
BC005743
LOCUS
DEFINITION
Mus musculus sideroflexin 1, mRNA (cDNA clone MGC:11926
IMAGE:3599636), complete cds.
ACCESSION
BC005743
VERSION
BC005743.1 GI:13543137
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

```

```

REFERENCE
AUTHORS
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 1204)
Strausberg,R.
12477932
2 (bases 1 to 1204)
Strausberg,R.
12477932

```

```

AUTHORS
DIRECT SUBMISSION
Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

```

FEATURES
source
1. .1204
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11926 IMAGE:3599636"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAK Plate: 16 Row: P Column: 6.
Location/Qualifiers


```

/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..1204
/gene="Sfxn1"
/note="synonym: 2810002005R1k"
/db_xref="LocusID:14057"
/db_xref="MGI:2137677"
57..1025
/codon_start=1
/product="sideroflexin 1"
/protein_id="AAH05743.1"
/db_xref="GI:13543138"
/db_xref="LocusID:14057"
/translation="MSGEVPPNINIKERPDQSTFIGRASHFTVTDPRNILLTNEQL
ENARKVVDYRQIVPAGLTENEELRAKYADSAFHPDGEKMTLIGMSAOVPMNMT
ITGCMTEFYRTTPAVLFWQWINSFNNAVNYNRSQDAPLVNELGTAYSATGAVA
TALGLNALTKRVSPILGRFVPAVAANANCINIPLMRQRELVGIPVTDENGTRLGES
TNAKQAITOVVISRIILMAAPGMAIPPFIMNTLEKKAFLKRPWMSAPIQVTLVGFCI
VFATPLCCALFPQKSSMSVTSLEDELQASIQRTHPAIRRYFNKGL"
BASE COUNT      307 a      309 c      322 g      266 t
ORIGIN

Query Match      35.1%; Score 432; DB 10; Length 1204;
Best Local Similarity 72.7%; Pred. No. 4.1e-87;
Matches 558; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY      465 GGTGAATGCGCTTTAGACATCAACATCCAGAACCTCGCTGGACCAAGTACTTCTCG 524
      11 111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1
Db      63 GGGGAAGTCCACCCCAACATTAACATCAAGAGACCTCGATGGACCAAGCACGTTCAATT 122

QY      525 GGCAGAGCCCGGCACCTTTTCACACTGTACTGATCCTCGAAATCTGCTGCTCGGGGCA 584
      111 1111 1 1111 1111 1111 1111 1111 1111 1111 1111 1111 1
Db      123 GGCCGAGCAGTCACTTCTTCACGGGTGATGATCCAGAAATATCTTTTAAAGAACGAA 182

QY      585 CAGCTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGCGCGCTGCTGACCCAGGG 644
      1111 11 1 1111 1111 1111 1111 1111 1111 1111 1111 1111 1
Db      183 CAGCTAGAGATGCGAGAAAGTGTGCTGACTACAGGCAAGAAATCGTTCCTGCGGCG 242

QY      645 ATCACCAGAGACCACTGTGAGAGGCCAAGTATGTATGACTCCGCCCTTCCATCCGAG 704
      1111 11 1 1111 1111 1111 1111 1111 1111 1111 1111 1111 1
Db      243 CTCACGGAATAGACTATGAGAGCGCAAGTACGCGTACGACTCGGCTTCCATCCTGAC 302

QY      705 ACAGGGAGAGAGGTGCTCTGATTTGCCCATGTCAGCCCAAGTGCCTCATGAACATGACC 764
      11 11 1111 11 1111 11 1111 11 1111 11 1111 11 1111 1111 1
Db      303 ACCGCTGAAGAATGACTCTGATAGAGAAATGTCAGCTCAGGTCGCCGATGAACATGACC 362

QY      765 ATCACTGGCTGATGCTCACAATTTCTACAGGAAGAACCCCAACCGTGTGTTCTGGCAGTGG 824
      1111 1111 1111 11 1111 111 1111 111 111 111 1111 1111 1111
Db      363 ATCACAGGCTGATGATGACCTTCTACCGGACCACACCGGCTGTGCTTTTCTGGCAGTGG 422

QY      825 GTGAATCACTCTTCAATGCCATGTTAACTACTCCAAACCGCAGTGTGACACTCCCATC 884
      1 1111 1111 1111 11 1111 1111 1111 1111 1111 1111 1111 11
Db      423 ATAAACCACTCTTCAATGCGGTGCTCACTACACCAACAGAAAGCGTGACGCTCCCTC 482

QY      885 ACTGTGAGCAGCTGGGACAGCCCTATGTGAGTGCACACCACTGAGCTGTGGCCACGGCC 944
      1111 1111 1111 11 1111 11 1111 1111 1111 1111 1111 1111
Db      483 ACTGTGAATGAGCTGGGAACGGCTTATGTTTCTGCAACCACTGGCGCTGTGGCAACGGCC 542

QY      945 CTGGACTCAAAATCCCTCACCAAGCACTGCCCCCTTGTGCGGAGATTGTGCCCTTT 1004
      11 1111 1111 11 1111 11 1111 11 1111 11 1111 1111 1111
Db      543 CTAGACTCAATGCTTAAACCAAGCGGTCTCCCCACTGATAGAGCGTTTGTCCCTTC 602

QY      1005 GCAGCAGTGGCAGCTGCCACTGCATCAACATCCCCCTGATGAGGACAGAGAGCTGCAG 1064
      11 11 11 1111 1111 1111 1111 1111 1111 11 11 1111 11
Db      603 GCTGCTGAGCTGCTGCTAACTGCTTAAACATCCCACTGATGAGACAAAGGAGCTGAAG 662

QY      1065 GTGGGCATCCCGGTGCTGATGAGGACGTCAGAGGCTTGGCTTACTCGGCTGACTGCAGCC 1124
      11 11 11 11 1111 11 1111 11 1111 11 11 1111 1111 1111
Db      663 GTTGTATTCTCTGTCACTGATGAATAATGGCACTCGCTTGGGAGAAATCGAACCAACGACGCA 722
```

```

QY      1125 AAGCAGGAATCTTCCAGGTGCTGATTTCAAGAAATCTGCATGGCGGATTCCTGCCATGGCC 1184
      1111 1 111 1111 1111 11 11 111 1111 11 1111 11
Db      723 AAGCAAGCCATCAGCAGAGGTGCTCATCTCCAGGATCTCATATGGACAGCCCGGCGCATGGCC 782

QY      1185 ATCCCAACCATGATCATGACACTCTGGAGAGAAAGACTTCCCTGAAG 1232
      1111 11 1111 1111 111 111 1111 111 111
Db      783 ATCCCTCGTTTATCATGAACACCTTGAAAGAAAGACCTTTTAAAG 830

RESULT 13
AF325260
LOCUS      2803 bp      mRNA      linear      ROD 25-APR-2001
DEFINITION Mus musculus sideroflexin 1 (Sfxn1) mRNA, complete cds; nuclear
gene for mitochondrial product.
ACCESSION AF325260
VERSION AF325260
KEYWORDS AF325260.1 GI:13785611
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2803)
AUTHORS Fleming,M.D., Campagna,D.R., Haslett,J.N., Trenor,C.C. III and
Andrews,N.C.
TITLE A mutation in a mitochondrial transmembrane protein is responsible
for the pleiotropic hematological and skeletal phenotype of
flexed-tail (f/f) mice
JOURNAL Genes Dev. 15 (6), 652-657 (2001)
MEDLINE 21172735
PUBMED 11274051
REFERENCE 2 (bases 1 to 2803)
AUTHORS Fleming,M.D., Campagna,D.R. and Andrews,N.C.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-2000) Pathology, Children's Hospital, 300
Longwood Avenue, Boston, MA 02115, USA
FEATURES
location/Qualifiers
1..2803
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="13"
1..2803
/gene="Sfxn1"
78..1046
/note="Sfxn1; mitochondrial membrane protein"
/codon_start=1
/product="sideroflexin 1"
/protein_id="AAK39428.1"
/db_xref="GI:13785612"
/translation="MSGEVPPNINIKERPDQSTFIGRASHFTVTDPRNILLTNEQL
ENARKVVDYRQIVPAGLTENEELRAKYADSAFHPDGEKMTLIGMSAOVPMNMT
ITGCMTEFYRTTPAVLFWQWINSFNNAVNYNRSQDAPLVNELGTAYSATGAVA
TALGLNALTKRVSPILGRFVPAVAANANCINIPLMRQRELVGIPVTDENGTRLGES
TNAKQAITOVVISRIILMAAPGMAIPPFIMNTLEKKAFLKRPWMSAPIQVTLVGFCI
VFATPLCCALFPQKSSMSVTSLEDELQASIQRTHPAIRRYFNKGL"
122..123
/gene="Sfxn1"
/note="insertion mutation results in the flexed-tail (f)
mouse"
variation
/replacement="aat"
BASE COUNT      784 a      644 c      646 g      729 t
ORIGIN

Query Match      35.1%; Score 432; DB 10; Length 2803;
Best Local Similarity 72.7%; Pred. No. 3.9e-87;
Matches 558; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY      465 GGTGAATGCGCTTTAGACATCAACATCCAGAACCTCGCTGGACCAAGTACTTCTCG 524
      11 111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1
Db      84 GGGGAAGTCCACCCCAACATTAACATCAAGAGCCTCGATGGACGACAGCAGTTCATT 143

QY      525 GGCAGAGCCCGGCACCTTTTTCACACTGTACTGATCCTCGAAATCTGCTGCTCGGGGCA 584
```

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REMARK
Db	144	GGCCGAGCCAGTCACTTCTTCACGGTTACTGATCCCGAATAATATCCTTTAAAGAACGAA	203									
QY	585	CAGCTGGAAGCTTCTCGGAACATCGTGACAGAACTACAGGGCCGGCTGGTGACCCCAAGG	644									
Db	204	CAGCTAGAGAATGCGAGGAAGTGGTGATGACTACAGGCCAAGAAATCGTTCCTGCGGCG	263									
QY	645	ATCACCGAGGACCAAGCTGTGAGGGCCAGATATGTATGACTCCGCCCTTCATCCGGAC	704									
Db	264	CTCAGCGAAAATGAGCTATGGAGAGCGAAGTACGGCTGACACTCGGCCCTTCATCTGAC	323									
QY	705	ACAGGGGAGAAGTGGTCTCTGATTGGCCCGCATGTCAAGCCCAAGGTGCCCATGAACATGACC	764									
Db	324	ACCGGTGAAAAGATGACTCTGATAGGAAGAATGTCAAGCTCAAGTCCGATGAACATGACC	383									
QY	765	ATCACTGGCTGCATGCTCACAATTCTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGG	824									
Db	384	ATCACAGGCTGCATGATGACCTTCTTACCGGACCAACCGGCTGTGCTTTCTGGCAGTGG	443									
QY	825	GTGAATCAGTCCCTCAATGCCATTGTTAACTACTACCAACCCAGTGTGACACATCCCATC	884									
Db	444	ATAAACCAAGTCCCTCAATGCCCGTGTGCTCAACTACACCAACAGAAGCGGTGACGCTCCCTC	503									
QY	885	ACTGTAGGACGAGTGGGGACAGCCCTATGTGAGTGGCCACCACCTGGAGCTGTGGCCACGGCC	944									
Db	504	ACTGTGAATGAGCTGGGGAACGGCTTATGTCTTCTGCAACCACTGGCGCTGTGGCAACGGCC	563									
QY	945	CTGGGACTCAAAATCCCTCACCAAGACACCTGCCCCCTTGGTCCGACAGATTGTGCCCTTT	1004									
Db	564	CTAGGACTCAATGCCCTTAAACCAAGCGTGTCTCCCACTGATAGGACGTTTGTCCCTTC	623									
QY	1005	GCAGCAGTGGCAGCTGCCAATGCAATGCAATCAACATCCCCCTGATGAGGACAGAGAGCTGCAG	1064									
Db	624	GCTGCTGTAGCTGTGCTTAAGTCAATTAACATCCCACTGATGAGACAAAGGAGCTGAAG	683									
QY	1065	GTGGGACATCCCGGTGGCTGATGAGGAGGTGACAGGCTTGGCTACTCGGTGATGACGACC	1124									
Db	684	GTTGTATTCCTGCTCACTGATGAAAATGGCACCTGCTTGGGAGAATCGACCAACGACGA	743									
QY	1125	AAGCAGGGAATCTTCCAGGTGGTGTGATTTCAAGAACTGATGCGCATTCCTGCCATGGCC	1184									
Db	744	AAGCAAGCCATCAGCAGAGTGTGATCTCCAGATCCTCATATGACGACCCCGGCGATGGCC	803									
QY	1185	ATCCCAACCACTGATCATGAGACACTCTGGAGAAGAAAGACTTCCCTGAAG	1232									
Db	804	ATCCCTCCGTTTATCATGAACACCTTGGAAGAAAGACCTTTTAAAG	851									
RESULT 14												
S70011		2986 bp mRNA linear ROD 10-JUL-2000										
LOCUS	S70011	Rattus sp. liver tricarboxylate carrier mRNA, partial cds; nuclear										
DEFINITION	S70011	gene for mitochondrial product.										
ACCESSION	S70011											
VERSION	S70011.1	GI:545997										
KEYWORDS												
SOURCE		Rattus sp.										
ORGANISM		Rattus sp.										
		Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;										
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;										
		Rattus.										
REFERENCE		1 (bases 1 to 2986)										
AUTHORS		Azzi, A., Glerum, M., Koller, R., Mertens, W. and Spycher, S.										
TITLE		The mitochondrial tricarboxylate carrier										
JOURNAL		J. Bioenerg. Biom										

		CDS	
		<pre> /mol_type="mRNA" /db_xref="taxon:10118" /tissue_type="liver" <1..1076 /note="This sequence comes from Fig. 9" /codon_start=3 /product="tricarboxylate carrier" /protein_id="AAB30258.1" /db_xref="GI:545998" /translation="APGGGARDRGASGSGRRHAGCPRGAGGGRGGSGTMSGEVPEPNTI NIKEPRWDSTFIIGRASHFTVTDPKNILLTMEQLENARKVVHDYRGQIVPAGLTENE LMRAKAYDSAFHPDGTGKMTLIGRMSAQVPMNMTITGCMATFYRTTPAVLFEQWINO SFNAVVNTYNRSGDAPLTVNEIGTAYVSATGAVATAGLNALTKHVSPILGRVPEFA AVAAANCINIPLMRQRELKVGIPVTDENGTRIGESTNAKAQAITQVVISRIILMAAPGM AIPFIMNTLEKKAFLKRFPMWSAPIQVTLVEGCLVFATPLCCALFPQKSSMSVTSLE DDLQASIOKSHPELRRVYFNKGL" </pre>	
BASE COUNT	818 a	690 c	697 g
ORIGIN			781 t
Query Match	34.38;	Score 422.4;	DB 10; Length 2986;
Best Local Similarity	71.9%;	Pred. No. 5.8e-85;	
Matches 552; Conservative	0;	Mismatches 216;	Indels 0; Gaps 0;

Query Match	34.3%	Score 422.4	DB 10	Length 2986
Best Local Similarity	71.9%	Pred. No. 5.8e-85		
Matches 552	Conservative 0	Mismatches 216	Indels 0	Gaps 0
QY 465	GGTGAATTCCTTTAGACATCAACATCCAGAACCTGCTGGGACCAAGTACTTTCC	524		
DB 114	GGGGAAGTGCACCCACCAACATTAAACATCAAGAGCCTCGATGGGACAGACAGCTTCAT	173		
QY 525	GGCAGAGCCCCGGCACTTTTCACTGTCTGATCCTCGAAATCTGCTGTCCGGGCA	584		
DB 174	GGCCGAGCCAGTCACTTCTTACGCGTTACTGATCCCAAAACATCCTTTTAACGAACGAA	233		
QY 585	CAGCTGGAAGCTTTCGGGAACATCGTGAGAACTACAGGGCCGGCTGTGACCCCAAGG	644		
DB 234	CAGCTAGAGATGCGAGGAAGTGTACACGATTAACGCAAGAAATCGTTCCTGCCGGC	293		
QY 645	ATCACCGAGGACCAGCTGTGGAGGGGCCAAGTATGTATGACTCCGCCCTTCATCCGAC	704		
DB 294	CTCACGGAAATGAGTTATGGAGAGCGGAAGTACCGCTATGACTCGGCCCTTCATCCGAC	353		
QY 705	ACAGGGGAGAAGGTGGTCTGATTGGCCGATGTACGCCAGGTGCCATGAACATGACC	764		
DB 354	ACGGGTGAAAAGATGACTCTGTATAGGAAGATGTGGCTCAGGTCGCCGATGAACATGACC	413		
QY 765	ATCACTGGCTGCATGCTCACAATTCTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGG	824		
DB 414	ATCACGGGCTGCATGATGACCTTCTACCGGACACACCGGCTGTGCTTCTTGCGAGTGG	473		
QY 825	GTGAATCAGTCCCTCAATGCCATTGTTAACTACTCCAACCGCAGTGGTGACATCCATC	884		
DB 474	ATAAACCAAGTCCCTCAATGCCGCTGTGCACTACACCAATAGAAGTGTGACGCTCCCTT	533		
QY 885	ACTGTAGGACAGCTGGGGACAGCCTATGTGAGTGGCAACCACTGGAGCTGTGGCCAGGCC	944		
DB 534	ACTGTAATGAGCTGGGAACGGCTTATGTTCTGCAACAACCGCGCGCTGGCGACAGCC	593		
QY 945	CTGGGACTCAAAATCCCTCAACCAAGCACCTGCCCCCTTTGGTCCGACAGATTGTGCCCTT	1004		
DB 594	CTTGGACTCAATGCCCTTAACCAAGCATGTGTGCGCGCTGATAGGACGTTTGTTCCTTC	653		
QY 1005	GCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGGACAGAGAGAGCTGCAG	1064		
DB 654	GCTGCTGATGCTGTGCTAACTGATTAACATCCCGCTGATGAGACAAAGGGAATCAAG	713		
QY 1065	GTGGCATCCCGGTGCTGATGAGGACAGTCAAGAGCTTGCTACTCGGTGACTGACGCC	1124		
DB 714	GTTGTTATTCCTGCTCACTGATGAGAAATGGACACAGCTGGGAGAAATCGAACCAAGCAGCA	773		
QY 1125	AAGCAGGAATCTTCAGGTGTGATTCAAGAACTGCATGGCGATTCTCGCATGACC	1184		
DB 774	AAGCAAGCCATCAAGCAGGTGTGATCTCCAGATCCTCATGCGCGCCCTGGCATGGCC	833		
QY 1185	ATCCCAACCACTGATCATGACACTCTGGAGAAAGAACTTCCTGAAG	1232		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 08:15:20 ; Search time 365 Seconds
(without alignments)
911.528 Million cell updates/sec

Title: US-09-990-415A-1
Perfect score: 1232
Sequence: 1 cccttagcgccagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1232	100.0	1232	24	ABN83754	Human protein clus
2	1231	99.9	1826	22	AAF27703	Human transport pr
c	1208.2	98.1	3047	22	AAH72722	Human cervical can
4	944	76.6	1729	22	AA159123	Human polynucleoti
5	880	71.4	1061	24	ABN83755	Human protein clus
6	869.4	70.6	1375	24	ABO98865	Human ORF672 codin
7	848.8	68.9	2702	22	AAK94713	Human full-length
8	818.6	66.4	984	24	AB139805	Human NS CDNA sequ

9	805.8	65.4	843	21	AAC75798	Human ORFX ORF1353
10	659	53.5	752	22	AAK91763	Human CDNA 5'-end
11	659	53.5	752	22	AAK93199	Human CDNA clone r
12	632.2	51.3	2215	23	AAS84362	DNA encoding novel
13	516.6	41.9	723	22	AA160909	Human polynucleoti
14	399.4	32.4	2269	24	ABN83757	Human protein clus
15	398.2	32.3	1076	22	AA159569	Human polynucleoti
16	398.2	32.3	1485	22	AA126555	Human breast cance
17	398.2	32.3	2053	22	AA161354	Human polynucleoti
18	398.2	32.3	2367	24	AAD21995	Human transporters
19	398.2	32.3	2468	21	AAC75794	Human ORFX ORF1349
20	396.8	32.2	1134	24	ABN84365	Alzheimer's diseas
21	396.6	32.2	2947	22	AAH44832	Rat CDNA encoding
22	385.4	31.3	1503	24	ABK54156	CDNA encoding huma
23	351.2	28.5	2766	22	AAH15610	Human CDNA sequenc
24	346.4	28.1	1084	24	ABK96715	DNA encoding human
25	335.4	27.2	1567	24	ABN83756	Human protein clus
26	332.2	27.0	1772	24	AAD33666	Human TRICH-21 CDN
27	332.2	27.0	2312	22	AA159647	Human polynucleoti
28	332.2	27.0	2376	22	AA157861	Human polynucleoti
29	330.6	26.8	2458	24	ABK35352	Human CDNA encodin
30	326.6	26.5	413	23	AAS84361	DNA encoding novel
31	315	25.6	984	24	ABK96716	DNA encoding splic
32	315	25.6	984	24	ABN84366	Alzheimer's diseas
33	302.8	24.6	2642	22	AAH17727	Human CDNA sequenc
34	302.4	24.5	1575	24	ABL99804	Human secretory po
35	288.4	23.4	741	22	AAH06089	Human CDNA clone (
36	278.6	22.6	1286	23	ABL18135	Drosophila melanog
37	276.8	22.5	1548	24	ABL90501	Human polynucleoti
38	243.8	19.8	507	23	ABV45979	Human prostate exp
39	232.2	18.8	421	23	ABV16181	Human prostate exp
40	225.6	18.3	524	22	AAS31167	Human diagnostic a
41	222.8	18.1	806	22	AAH04784	Human CDNA clone (
c	173.6	14.1	4438	23	ABL11786	Drosophila melanog
43	170.6	13.8	1585	23	ABL11787	Drosophila melanog
44	164.2	13.3	1244	20	AAK97798	Extended human sec
45	162	13.1	2673	24	AAD36304	Human transporter

ALIGNMENTS

RESULT 1	
ABN83754	
ID	ABN83754 standard; CDNA; 1232 BP.
XX	
AC	ABN83754;
XX	
DT	02-SEP-2002 (first entry)
XX	
DE	Human protein cluster I gene.
XX	
KW	Protein cluster I; human; metabolic disorder; obesity; diabetes;
KW	antidiabetic; diagnosis; therapy; chromosome 10; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	CDS
FT	
FT	location/Qualifiers
FT	450..1232
FT	/*tag= a
FT	/product= "Protein cluster I"
FT	/partial
FT	/note= "the CDS does not include a stop codon"
XX	
PN	WO200242324-A1.
XX	
PD	30-MAY-2002.
XX	
PF	22-NOV-2001; 2001WO-SE02581.
XX	
PR	24-NOV-2000; 2000SE-0004325.
XX	
PA	(PHAA) PHARMACIA AB.

XX
PI
XX
DR
DR

WPI; 2002-500277/53.
P-PSDB; ABB76445.

Altersand A;

PT Novel nucleic acid molecule encoding Protein Cluster I, useful in the
PT diagnosis of metabolic diseases, such as obesity and diabetes, and in
PT the identification of agents useful in the treatment of the diseases
XX
PS Claim 1; Page 23-24; 34pp; English.

Claim 1; Page 23-24; 34pp; English.

The present invention relates to the identification of a human gene family (see ABN83754-57) encoding a group of polypeptides (see ABB76445-48) referred to as Protein Cluster I. This family of homologous proteins was identified by an 'all-versus-all' BLAST procedure using all *Caenorhabditis elegans* proteins in the Wormpep20 database release. Proteins of unknown function were compared to the *Drosophila melanogaster* Flybase database, and non-annotated protein clusters, conserved in both *C. elegans* and *D. melanogaster*, were used in a BLAST procedure against the *Celera* Human Genome Database, and Protein Cluster I proteins of unknown function were selected for study. The human part of protein Cluster I comprises polypeptides encoded by 3 genes; an alternative splicing of the present sequence (deletion of nucleotides 624-794) gives the sequence in ABN83755. The present gene was identified in a human DNA sequence from clone RP11-108L7 on chromosome 10. It is expressed primarily in the nervous system and digestive system. The claimed Protein Cluster I nucleic acid molecules and proteins are proposed to be useful for differential identification of the tissue(s) or cell type(s) present in a biological sample, for diagnosis of diseases and disorders, including metabolic disorders and immune disorders, especially obesity and diabetes, and for identifying agents useful in the treatment of such diseases. The nucleic acid molecules are also useful as hybridisation probes, for chromosome and gene mapping, in PCR technology, and in the production of sense or antisense nucleic acids.

SQ Sequence 1232 BP; 239 A; 368 C; 375 G; 250 T; 0 other;

Query Match	100.0%;	Score 1232;	DB 24;	Length 1232;
Best Local Similarity	100.0%;	Pred. No. 1.3e-303;		
Matches 1232; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0

QY	1	CCCTTAGGCGCCAGGACAGCCGAGCGTTACCTGCTCCGGGACGCGAGTTCCTTAAACC	60
Db	1	CCCTTAGGCGCCAGGACAGCCGAGCGTTACCTGCTCCGGGACGCGAGTTCCTTAAACC	60
QY	61	ACCCAGTCTGTGTTCTGACGCGCTAGCTCATTCGCAATTTAGGGCTGGGTCGTGGCT	120
Db	61	ACCCAGTCTGTGTTCTGACGCGCTAGCTCATTCGCAATTTAGGGCTGGGTCGTGGCT	120
QY	121	TGTTCCCTCCGGCTCGAACCACCTCTTCTCTGAGCCGAGCCAGCTACCGGGGCTCCTGG	180
Db	121	TGTTCCCTCCGGCTCGAACCACCTCTTCTCTGAGCCGAGCCAGCTACCGGGGCTCCTGG	180
QY	181	AATTGCCAACCCTCCCTGGGACACCTTGAGGCCCTCCGTGAGGAGGACGTACGGGGGACAG	240
Db	181	AATTGCCAACCCTCCCTGGGACACCTTGAGGCCCTCCGTGAGGAGGACGTACGGGGGACAG	240
QY	241	CGGAGCGTGAGCCTGAGTTTGTCTGACAGGCGTGCTGTGTGTGGCTGGGTTCTGCCAAT	300
Db	241	CGGAGCGTGAGCCTGAGTTTGTCTGACAGGCGTGCTGTGTGTGGCTGGGTTCTGCCAAT	300
QY	301	CCCCGTGCCCCACCGGAGTGGGCGCGCGCGGGAAGCTCCTGCCCCCTCCCTGCTGGTCGGCGT	360
Db	301	CCCCGTGCCCCACCGGAGTGGGCGCGCGCGGGAAGCTCCTGCCCCCTCCCTGCTGGTCGGCGT	360
QY	361	CACGCGTGACGTCCCGCGGTGATGGCTGGGAGGGGCCCGGCGGACAGCGGAGCAGAGAG	420
Db	361	CACGCGTGACGTCCCGCGGTGATGGCTGGGAGGGGCCCGGCGGACAGCGGAGCAGAGAG	420
QY	421	GAAAGCGGTTCTGAGAGCTTCAGAGCGATGGAAGCAAAATGGGTGAATTGCTTTAG	480

[illegible]

XX 28-DEC-2000.
PD
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzal Y, Lu DAM, Au-Young J, Patterson C;
XX
DR WPI: 2001-041424/05.
DR P-PSDB; AAB60083.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 5; Page 142-143; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTS). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 1826 BP; 415 A; 519 C; 512 G; 380 T; 0 other;

Query Match 99.9%; Score 1231; DB 22; Length 1826;
Best Local Similarity 100.0%; Pred. No. 2.6e-303;
Matches 1231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCTTAGGCGCCAGGAGACCGCAGCTTACCTGTCCCGGCGAGCGAGTTCTTTACCA 61
Db 1 CCTTAGGCGCCAGGAGACCGCAGCTTACCTGTCCCGGCGAGCGAGTTCTTTACCA 60
QY 62 CCCAGTTCTGTTCTGACGCGCCCTAGCTATTCCGCAAAATTAGGGCTTGGTCTGGCTT 121
Db 61 CCCAGTTCTGTTCTGACGCGCCCTAGCTATTCCGCAAAATTAGGGCTTGGTCTGGCTT 120
QY 122 GTTCCCTCCGGCTCGAACCACTTCTCTGAGCCGAGCCAGCTACCGGGCTCTTGA 181
Db 121 GTTCCCTCCGGCTCGAACCACTTCTCTGAGCCGAGCCAGCTACCGGGCTCTTGA 180
QY 182 ATGCGCACCCCTCCCTGGGACCCCTTGAGGCTCCCTGAGGAGCTACGGGGCAGAGC 241
Db 181 ATGCGCACCCCTCCCTGGGACCCCTTGAGGCTCCCTGAGGAGCTACGGGGCAGAGC 240
QY 242 GGGAGCTGAGCCTGAGTTGCTGAGGCGCTGCTGTGTGGTGGCTGGTCTGCCAATC 301
Db 241 GGGAGCTGAGCCTGAGTTGCTGAGGCGCTGCTGTGTGGTGGCTGGTCTGCCAATC 300
QY 302 CCCGTGCCACCGGGTGGGCGCGCGGGAAGCTCCTGCCCTCCTGCTGTGGCGCTC 361
Db 301 CCCGTGCCACCGGGTGGGCGCGCGGGAAGCTCCTGCCCTCCTGCTGTGGCGCTC 360
QY 362 ACGCGTGACGTCCCGCTGATGGCTGGGAGGCGCGGCGGAGACGCGGAGAGAGG 421
Db 361 ACGCGTGACGTCCCGCTGATGGCTGGGAGGCGCGGCGGAGACGCGGAGAGAGG 420
QY 422 AAGCGGTTCTGAGAGCTTCAAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTGA 481
Db 421 AAGCGGTTCTGAGAGCTTCAAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTGA 480
QY 482 CATCAACATCCAGAACCTCGCTGGGAGCAAAAGTACTTCTTGGGAGAGAGCGCGCACTT 541
Db 481 CATCAACATCCAGAACCTCGCTGGGAGCAAAAGTACTTCTTGGGAGAGAGCGCGCACTT 540
QY 542 TTTCACGTGTACTGATCTCTGAAATCTGCTGTGTCCGGGCGACAGCTGGAAGCTTCTCG 601

Db 541 TTTCACGTGTACTGATCTCTGAAATCTGCTGTGTCCGGGCGACAGCTGGAAGCTTCTCG 600
QY 602 GAACATCGTGACAACTACAGGGCGCGGTGTGACCCCAAGGATCACCGAGACAGCT 661
Db 601 GAACATCGTGACAACTACAGGGCGCGGTGTGACCCCAAGGATCACCGAGACAGCT 660
QY 662 GTGAGGGGCAAGTATGTGTATGACTCCGCTTCCATCCGGACACAGGGAGAAAGTGT 721
Db 661 GTGAGGGGCAAGTATGTGTATGACTCCGCTTCCATCCGGACACAGGGAGAAAGTGT 720
QY 722 CCTGATTGGCCGATGTGAGCCAGGTCGCCATGAACATGACCATGAGCTGCTGATGCT 781
Db 721 CCTGATTGGCCGATGTGAGCCAGGTCGCCATGAACATGACCATGAGCTGCTGATGCT 780
QY 782 CACATTTACAGGAAGACCCCAACCGTGTGTCTGGCAGTGGGTGAATCAGTCTTCAA 841
Db 781 CACATTTACAGGAAGACCCCAACCGTGTGTCTGGCAGTGGGTGAATCAGTCTTCAA 840
QY 842 TGCCATTGTTAACTACTCCAAACCGCAGTGTGACACTCCCATCATGTGAGGAGCTGGG 901
Db 841 TGCCATTGTTAACTACTCCAAACCGCAGTGTGACACTCCCATCATGTGAGGAGCTGGG 900
QY 902 GACAGCCTATGTGAGTGCCACCACTGAGCTGTGGCCACGCGCCCTGGAGCTCAATCCCT 961
Db 901 GACAGCCTATGTGAGTGCCACCACTGAGCTGTGGCCACGCGCCCTGGAGCTCAATCCCT 960
QY 962 CACCAAGCACTGCCCCCTTGGTGGCAGATTGTGCCCCCTTGGCAGCAGTGCGAGCTGC 1021
Db 961 CACCAAGCACTGCCCCCTTGGTGGCAGATTGTGCCCCCTTGGCAGCAGTGCGAGCTGC 1020
QY 1022 CAACTGCATCAACATCCCCCTGATGAGCGAGAGAGAGCTGACGTGGGCTATCCCGGTGGC 1081
Db 1021 CAACTGCATCAACATCCCCCTGATGAGCGAGAGAGAGCTGACGTGGGCTATCCCGGTGGC 1080
QY 1082 TGATGAGCAGGTTCAGAGGCTTGGCTACTCGGTGACTGACGCCAAGCAAGGAATCTTCA 1141
Db 1081 TGATGAGCAGGTTCAGAGGCTTGGCTACTCGGTGACTGACGCCAAGCAAGGAATCTTCA 1140
QY 1142 GGTGGATTTCAAGAAATCTGCATGGCGATTCCTGCGCATGGCCATCCACCATGATCAT 1201
Db 1141 GGTGGATTTCAAGAAATCTGCATGGCGATTCCTGCGCATGGCCATCCACCATGATCAT 1200
QY 1202 GGACACTCTGAGAAAGAAAGACTTCTGAAG 1232
Db 1201 GGACACTCTGAGAAAGAAAGACTTCTGAAG 1231

RESULT 3
AAH72722/c
ID AAH72722 standard; cDNA; 3047 BP.
XX
AC AAH72722;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3996.
XX
KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.

```
PR 21-JUL-2000; 2000US-0220114.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX PT and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 815-816; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded
XX CC polypeptides are useful: to assess if a patient is afflicted with
XX CC cervical cancer or has a pre-malignant condition; to monitor the
XX CC progression of cervical cancer or a premalignant condition in a patient;
XX CC and to select and/or assess the efficacy of a compound or therapy for
XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX CC useful for gene therapy.
XX
XX Sequence 3047 BP; 629 A; 790 C; 891 G; 733 T; 4 other;
SQ

Query Match          98.1%; Score 1208.2; DB 22; Length 3047;
Best Local Similarity 99.8%; Pred. No. 2e-297;
Matches 1210; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 20 GCCGAGCGTTACCTGTGTCGCCGGGACGGAGTCTTTACCCACCCAGTCTGGTCTGA 79
DB 3035 GTCCGGCGTTACCTGTGTCGCCGGGACGGAGTCTTTACCCACCCAGTCTGGTCTGA 2976
QY 80 CGCCCTAGCTCATTCGCCAATTTAGGGCTTGGGCTGGGCTTCTCCCTCCGGCTCGAA 139
DB 2975 CGCCCTAGCTCATTCGCCAATTTAGGGCTTGGGCTGGGCTTCTCCCTCCGGCTCGAA 2916
QY 140 CCACCTCTTCTGTAGAGCCGAGCAGTACCGGGGCTCCTGGAATTGCCACCCCTCCCTGG 199
DB 2915 CCACCTCTTCTGTAGAGCCGAGCAGTACCGGGGCTCCTGGAATTGCCACCCCTCCCTGG 2856
QY 200 GCACCTTGAGGCTCCGTGGAGGAGCTCACGGGGGACAGCGGAGCGTGAAGCTGAGTT 259
DB 2855 GCACCTTGAGGCTCCGTGGAGGAGCTCACGGGGGACAGCGGAGCGTGAAGCTGAGTT 2796
QY 260 TGTGTCAGGCGTGTCTGTGTGGGCTGGGCTTCTGCCAATCCCGGTGCCCCACCGGGTGG 319
DB 2795 TGTGTCAGGCGTGTCTGTGTGGGCTGGGCTTCTGCCAATCCCGGTGCCCCACCGGGTGG 2736
QY 320 GCGCGGCGCGGGAAGCTCTGCCCCCTCCCTGTGTGTCGGCGTCACGGCGTGACGTCCCGCT 379
DB 2735 GCGCGGCGCGGGAAGCTCTGCCCCCTCCCTGTGTGTCGGCGTCACGGCGTGACGTCCCGCT 2676
QY 380 GATGGCTGGGAGGGCGCGCGGACAGCGGAGGAGGAGAGAGAGGCGGTTCTGAGAGCT 439
DB 2675 GATGGCTGGGAGGGCGCGCGGACAGCGGAGGAGGAGAGAGGCGGTTCTGAGAGCT 2616
QY 440 TCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTTAGACATCAACATCCAGGAACC 499
DB 2615 TCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTTAGACATCAACATCCAGGAACC 2556
QY 500 TCGCTGGGACCAAAAGTACTTCTCTGGGACAGAGCCCGGACATTTTCACTGTACTGATCC 559
DB 2555 TCGCTGGGACCAAAAGTACTTCTCTGGGACAGAGCCCGGACATTTTCACTGTACTGATCC 2496
QY 560 TCGAAATCTGCTGTCTCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGAGAACTA 619
DB 2495 TCGAAATCTGCTGTCTCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGAGAACTA 2436
QY 620 CAGGGCCGGCGTGTGACCCCAAGGATCACCGAGACAGCTGTGAGGGGCCAAGTATGT 679
DB 2435 CAGGGCCGGCGTGTGACCCCAAGGATCACCGAGACAGCTGTGAGGGGCCAAGTATGT 2376
QY 680 GTATGACTCCGCTTCCATCCGAGACACAGGGGAGAAGTGTCTGATTGGCCCGCATGTC 739
```

```
DB 2375 GTATGACTCCGCTTCCATCCGAGACAGGGGAGAAGTGTCTGATTGGCCGATGTC 2316
QY 740 AGCCAGGTGCCCATGAACATGACCATCATGCTCATGCTCATATTCTACAGGAAGAC 799
DB 2315 AGCCAGGTGCCCATGAACATGACCATCATGCTCATGCTCATATTCTACAGGAAGAC 2256
QY 800 CCAACCGTGTGTCTTGCGCAGTGGGTAATCAGTCCCTCAATGCCATGTACTACTC 859
DB 2255 CCAACCGTGTGTCTTGCGCAGTGGGTAATCAGTCCCTCAATGCCATGTACTACTC 2196
QY 860 CAACCGCAGTGTGACACTCCCATCACTGTGAGGCAAGCTGGGACAGCCTATGTAGTGC 919
DB 2195 CAACCGCAGTGTGACACTCCCATCACTGTGAGGCAAGCTGGGACAGCCTATGTAGTGC 2136
QY 920 CACCACCTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACTGCCCCC 979
DB 2135 CACCACCTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACTGCCCCC 2076
QY 980 CTGTGTCGCGAGATTGTGCCCCCTTTCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCC 1039
DB 2075 CTGTGTCGCGAGATTGTGCCCCCTTTCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCC 2016
QY 1040 CCTGATGAGGACAGAGAGAGCTGACAGTGGGATCCCGGTGCTGATGAGGACAGTCAAG 1099
DB 2015 CCTGATGAGGACAGAGAGAGCTGACAGTGGGATCCCGGTGCTGATGAGGACAGTCAAG 1956
QY 1160 CTGCATGGCGATTCCTGCCATGGCCATGCCACCACTGATCATGACACTCTGAGAGAAG 1219
DB 1895 CTGCATGGCGATTCCTGCCATGGCCATGCCACCACTGATCATGACACTCTGAGAGAAG 1836
QY 1220 AGACTTCTCTGAAG 1232
DB 1835 AGACTTCTCTGAAG 1823

RESULT 4
AA159123
ID AA159123 standard; cDNA; 1729 BP.
XX
AC AA159123;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1326.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
```


XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR MPI: 2001-442253/47.
DR P-PSDB: AAM39967.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1326; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1729 BP; 406 A; 469 C; 491 G; 363 T; 0 other;

Query Match 76.6%; Score 944; DB 22; Length 1729;
Best Local Similarity 100.0%; Pred. No. 3e-230;
Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GGTCTGCAATCCCGTCCGCGGGGTGGCGCGCGCGGGAAGCTCCTCCCTCCT 348
Db 86 GGTCTGCCAATCCCGTCCGCGGGGTGGCGCGCGCGGGAAGCTCCTCCCTCCT 145
QY 349 GCTGCTCGCGCTCAGCGCGTGAAGTCCCGCGTGTGCTGGAGAGGCGCGCGCGGACAC 408
Db 146 GCTGCTCGCGCTCAGCGCGTGAAGTCCCGCGTGTGCTGGAGAGGCGCGCGCGGACAC 205
QY 409 GGAGGCAAGAGGAGGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 468
Db 206 GGAGGCAAGAGGAGGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 265
QY 469 AATTGCTTTAGACATCAACATCCAGAACCTCGCTGGACCAAAATGCTTCTGGGCA 528
Db 266 AATTGCTTTAGACATCAACATCCAGAACCTCGCTGGACCAAAATGCTTCTGGGCA 325
QY 529 GAGCCCGGCACTTTTCACTGTTACTGATCCTCGAATCTGCTGCTGCCGGGACACAC 588
Db 326 GAGCCCGGCACTTTTCACTGTTACTGATCCTCGAATCTGCTGCTGCCGGGACACAC 385
QY 589 TCGAAGCTTCTCGAACAATCGTCAGAACTACAGAGGCGCGGTGTGACCCAGGATCA 648
Db 386 TCGAAGCTTCTCGAACAATCGTCAGAACTACAGAGGCGCGGTGTGACCCAGGATCA 445
QY 649 CCGAGGACCAAGCTGTGAGGCGCAAGTATGTATGATTCGCCCTTCATCCGGACACAG 708
Db 446 CCGAGGACCAAGCTGTGAGGCGCAAGTATGTATGATTCGCCCTTCATCCGGACACAG 505
QY 709 GGGAGAAGGTGCTGATGTTGGCGCATGTACGCCAGGTGCCATGAACATGACCATCA 768
Db 506 GGGAGAAGGTGCTGATGTTGGCGCATGTACGCCAGGTGCCATGAACATGACCATCA 565
QY 769 CTGGCTGCATGCTACATCTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGGTGA 828
Db 566 CTGGCTGCATGCTACATCTACAGGAAGACCCCAACCGTGTGTTCTGGCAGTGGTGA 625

QY 829 ATCAGTCTTCAATGCGCATGTGTTAACTACTCCAAACCGCAGTGGTGACACTCCCATCACTG 888
Db 626 ATCAGTCTTCAATGCGCATGTGTTAACTACTCCAAACCGCAGTGGTGACACTCCCATCACTG 885
QY 889 TGAGGCAAGCTGGGGACAGCGCTATGTGAGTGGCCACCACCTGGAGCTGTGGCCACGCGCTGG 948
Db 686 TGAGGCAAGCTGGGGACAGCGCTATGTGAGTGGCCACCACCTGGAGCTGTGGCCACGCGCTGG 745
QY 949 GACTCAATCCCTCACCACCAAGCAGCTGCCCCCTTGGTGGGAGATTTGTGCCCTTTGCA 1008
Db 746 GACTCAATCCCTCACCACCAAGCAGCTGCCCCCTTGGTGGGAGATTTGTGCCCTTTGCA 805
QY 1009 CAGTGGCAGCTGCCAAGCTGCATCAACATCCCTGATGAGGAGAGAGCTGCAGGTGG 1068
Db 806 CAGTGGCAGCTGCCAAGCTGCATCAACATCCCTGATGAGGAGAGAGCTGCAGGTGG 865
QY 1069 GCAATCCCGTGGCTGATGAGGAGGCTCAGAGGCTTGGCTACTCGGTGACTGCAGCCAAAGC 1128
Db 866 GCAATCCCGTGGCTGATGAGGAGGCTCAGAGGCTTGGCTACTCGGTGACTGCAGCCAAAGC 925
QY 1129 AGGGAATCTTCCAGGTGTTGATTTCAAGATCTGCATGGCGAATTCCTGCCATGGCCATCC 1188
Db 926 AGGGAATCTTCCAGGTGTTGATTTCAAGATCTGCATGGCGAATTCCTGCCATGGCCATCC 985
QY 1189 CACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCCCTGAAG 1232
Db 986 CACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCCCTGAAG 1029

RESULT 5
ABN83755
ID ABN83755 standard; cDNA; 1061 BP.
XX
AC ABN83755;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human protein cluster I gene.
XX
KW Protein cluster I; human; metabolic disorder; obesity; diabetes;
KW antidiabetic; diagnosis; therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 450..680
FT FT /*tag= a
FT FT /product= "Protein cluster I"
FT FT /partial
FT FT /note= "the CDS does not include a stop codon"
XX
PN WO200242324-A1.
XX
PD 30-MAY-2002.
XX
PF 22-NOV-2001; 2001WO-SE02581.
XX
PR 24-NOV-2000; 2000SE-0004325.
XX
PA (PHAA) PHARMACIA AB.
XX
PI Altersand A;
XX
XX MPI: 2002-500277/53.
DR P-PSDB: ABB76446.
XX
XX Novel nucleic acid molecule encoding Protein Cluster I, useful in the
XX diagnosis of metabolic diseases, such as obesity and diabetes, and in
XX the identification of agents useful in the treatment of the diseases
PS Claim 1; Page 25-26; 34pp; English.
XX

CC The present invention relates to the identification of a human gene
CC family (see ABN83754-57) encoding a group of polypeptides (see
CC ABB76445-48) referred to as Protein Cluster I. This family of
CC homologous proteins was identified by an 'all-versus-all' BLAST
CC procedure using all Caenorhabditis elegans proteins in the
CC Wormpep20 database release. Proteins of unknown function were
CC compared to the Drosophila melanogaster Flybase database, and
CC non-annotated protein clusters, conserved in both C. elegans and D.
CC melanogaster, were used in a BLAST procedure against the Celera
CC Human Genome Database, and Protein Cluster I proteins of unknown
CC function were selected for study. The human part of Protein
CC Cluster I comprises polypeptides encoded by 3 genes. The present
CC sequence is an alternatively spliced version of the sequence given
CC in ABN83754. It is expressed primarily in the male genitalia.
CC The claimed Protein Cluster I nucleic acid molecules and proteins
CC are proposed to be useful for differential identification of the
CC tissue(s) or cell type(s) present in a biological sample, for
CC diagnosis of diseases and disorders, including metabolic disorders
CC and immune disorders, especially obesity and diabetes, and for
CC identifying agents useful in the treatment of such diseases. The
CC nucleic acid molecules are also useful as hybridisation probes, for
CC chromosome and gene mapping, in PCR technologies, and in the
CC production of sense or antisense nucleic acids.

XX Sequence 1061 BP; 203 A; 318 C; 322 G; 218 T; 0 other;

Query Match 71.4%; Score 880; DB 24; Length 1061;
Best Local Similarity 86.1%; Pred. No. 5e-214;
Matches 1061; Conservative 0; Mismatches 0; Indels 171; Gaps 1;

QY 1 CCCTTAGGCGCCAGGACACCGGAGCGTTACCTGGTCCGGGCGACGGAGTCTTTACCC 60
Db 1 CCCTTAGGCGCCAGGAGACGCGGAGCGTTACCTGGTCCGGGCGACGGAGTCTTTACCC 60
QY 61 ACCCGAGTCTGTTCTGAGCGCCCTAGCTCATTCGCAATTGAGGCTTGCGTGGCT 120
Db 61 ACCCGAGTCTGTTCTGAGCGCCCTAGCTCATTCGCAATTGAGGCTTGCGTGGCT 120
QY 121 TGTTCCTCCGCGCTGGAACCACTCTTCTCTGAGCCGAGCCAGCTACCGGGCTCTGG 180
Db 121 TGTTCCTCCGCGCTGGAACCACTCTTCTCTGAGCCGAGCCAGCTACCGGGCTCTGG 180
QY 181 AATTGCCACCCCTCCCTGGGACCCCTTGAGGCTCCGCTGGAGGAGCTCACGGGGCAGAG 240
Db 181 AATTGCCACCCCTCCCTGGGACCCCTTGAGGCTCCGCTGGAGGAGCTCACGGGGCAGAG 240
QY 241 CGGGACGTGAGCCTGAGTTGCTGCAGGCGTGTCTGTGTGGTGGCTGGTCTGCCAAT 300
Db 241 CGGGACGTGAGCCTGAGTTGCTGCAGGCGTGTCTGTGTGGTGGCTGGTCTGCCAAT 300
QY 301 CCCCGTGGCCACCGGTTGGGCGCGGGGGAAGCTCTGCCCCCTCCTGCTGTGGCGCT 360
Db 301 CCCCGTGGCCACCGGTTGGGCGCGGGGGAAGCTCTGCCCCCTCCTGCTGTGGCGCT 360
QY 361 CACGCGTGACGTCCCGCGTGATGGCTGGAGGGGCGCGGCGGACAGCGGAGGACAGAG 420
Db 361 CACGCGTGACGTCCCGCGTGATGGCTGGAGGGGCGCGGCGGACAGCGGAGGACAGAG 420
QY 421 GAAGGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTGCCCTTAG 480
Db 421 GAAGGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTGCCCTTAG 480
QY 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCTGGGACAGAGCCGGCACT 540
Db 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCTGGGACAGAGCCGGCACT 540
QY 541 TTTTCACTGTACTGATCCTCGAATATGCTGTGCTCCGGGGACACAGCTGAAGCTTCTC 600
Db 541 TTTTCACTGTACTGATCCTCGAATATGCTGTGCTGCTCCGGGGACACAGCTGAAGCTTCTC 600
QY 601 GGAACATCGTGACAGACTACAGGGCCGGCGTGGTGACCCAGGGATCACCGAGGACACAC 660
Db 601 GGAACATCGTGACAGAA----- 616

QY 661 TGTGGAGGGCCAGATATGTATGACTCCGCCCTTCATCCGGACACAGGGAGAAGTGG 720
Db 617 ----- 616
QY 721 TCCTGATTGGCCGCGATGTACGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGC 780
Db 617 ----- 616
QY 781 TCACATTTCTACAGGAAGACCCCAACCGTGCTGTCTGCGAGTGGGTGAATCAGTCCTTCA 840
Db 617 -----CTACAGGAAGACCCCAACCGTGCTGTCTGCGAGTGGGTGAATCAGTCCTTCA 669
QY 841 ATGCCATTGTTAACTACTTCAACCGCAGTGTGACACTTCCATCACTGTGAGGACAGCTGG 900
Db 670 ATGCCATTGTTAACTACTTCAACCGCAGTGTGACACTTCCATCACTGTGAGGACAGCTGG 729
QY 901 GGACAGCCTATGTAGTGCACCACTGAGCTGTGGCCAGGCCCTGGACTCAAAATCCC 960
Db 730 GGACAGCCTATGTAGTGCACCACTGAGCTGTGGCCAGGCCCTGGACTCAAAATCCC 789
QY 961 TCACCAAGCACCTGCCCCCTTGGTGGCAGATTGTGCCCTTTGACAGCAGTGGCAGCTG 1020
Db 790 TCACCAAGCACCTGCCCCCTTGGTGGCAGATTGTGCCCTTTGACAGCAGTGGCAGCTG 849
QY 1021 CCACTGCATCAACATCCCCCTGATGAGGACAGAGAGAGCTGCAGGTGGCATCCCGGTGG 1080
Db 850 CCACTGCATCAACATCCCCCTGATGAGGACAGAGAGAGCTGCAGGTGGCATCCCGGTGG 909
QY 1081 CTGATGAGGCGAGTCAAGAGGCTTGCGTACTCGGTGACTGCACCAAGCAGGGAATCTTCC 1140
Db 910 CTGATGAGGCGAGTCAAGAGGCTTGCGTACTCGGTGACTGCACCAAGCAGGGAATCTTCC 969
QY 1141 AGGTGTTGATTTCAGAAATCTGCATGGCGATTCTCGCCATGGCCATCCACACCTGATCA 1200
Db 970 AGGTGTTGATTTCAGAAATCTGCATGGCGATTCTCGCCATGGCCATCCACACCTGATCA 1029
QY 1201 TGGACACTCTGGAGAAGAAAGACTTCTGAAAG 1232
Db 1030 TGGACACTCTGGAGAAGAAAGACTTCTGAAAG 1061

RESULT 6
ABQ98865
ID ABQ98865 standard; DNA; 1375 BP.
XX
AC ABQ98865;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF672 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-0867550.
XX
PR 30-MAY-2000; 2000US-208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX WPI: 2002-626554/67.
DR P-PSDB: ABP64302.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease -
XX
PS Claim 2; SEQ ID 1343; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/sequence.html?DocID=20020082206.
XX
SQ Sequence 1375 BP; 342 A; 371 C; 374 G; 286 T; 2 other;

Query Match . 70.6%; Score 869.4; DB 24; Length 1375;
Best Local Similarity . 99.9%; Pred. No. 2.7e-211;
Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 362 ACCGCTGACGTCGCCGCTGATGCTGGAGGGCCCGCGGACAGCGGAGGACAGAGG 421
DB 3 ACGCGTGACGTCGCCGCTGATGCTGGAGGGCCCGCGGACAGCGGAGGACAGAGG 62
OY 422 AAGCGGCTTCTGAGAGCTTCAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTGA 481
DB 63 AAGCGGCTTCTGAGAGCTTCAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTGA 122
OY 482 CATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCCCTGGGACAGAGCGGCACTT 541
DB 123 CATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCCCTGGGACAGAGCGGCACTT 182
OY 542 TTTCACCTGTACTGATCCTCGAAATCTGCTGCTCCGGGGACACAGCTGGAAGCTTCTCG 601
DB 183 TTTCACCTGTACTGATCCTCGAAATCTGCTGCTCCGGGGACACAGCTGGAAGCTTCTCG 242
OY 602 GAACATCGTGCAAGACTACAGAGCGCGGCTGGTGAGACCCCAAGGATCACCGAGGACAGCT 661
DB 243 GAACATCGTGCAAGACTACAGAGCGCGGCTGGTGAGACCCCAAGGATCACCGAGGACAGCT 302
OY 662 GTGGAGGGCCAAAGTATGTATGATCTCCGCTTCATCCGGACACAGGGAGAGAGGTGCT 721
DB 303 GTGGAGGGCCAAAGTATGTATGATCTCCGCTTCATCCGGACACAGGGAGAGAGGTGCT 362
OY 722 CCTGATTGGCCGATGTACAGCCAGGTGCCATGAACATGACCATCACTGGCTGCATGCT 781
DB 363 CCTGATTGGCCGATGTACAGCCAGGTGCCATGAACATGACCATCACTGGCTGCATGCT 422
OY 782 CACATTTCTACAGGAAGACCCCAACCGTGGTGTCTGCGAGTGGTGAATCAGTCTTCAA 841
DB 423 CACATTTCTACAGGAAGACCCCAACCGTGGTGTCTGCGAGTGGTGAATCAGTCTTCAA 482
OY 842 TGGCATTTGTTAACTACTCCAAACCGCAGTGTGACACTCCCATCACTGTGAGGACAGCTGG 901
DB 483 TGGCATTTGTTAACTACTCCAAACCGCAGTGTGACACTCCCATCACTGTGAGGACAGCTGG 542
OY 902 GACAGCCTATGTGAGTGCACCACTGAGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCT 961
DB 543 GACAGCCTATGTGAGTGCACCACTGAGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCT 602

OY 962 CACCAAGACACTGCCCCCTTGGTGGCAGAGATTGTGCCCTTGGACAGAGTGACAGCTGC 1021
DB 603 CACCAAGACACTGCCCCCTTGGTGGCAGAGATTGTGTGCCCTTGGACAGAGTGACAGCTGC 662
OY 1022 CAATGCATCAACATCCCCCTGATGAGGCAGAGAGAGAGCTGCAGGTGGCATCCCGGTGC 1081
DB 663 CAATGCATCAACATCCCCCTGATGAGGCAGAGAGAGAGCTGCAGGTGGCATCCCGGTGC 722
OY 1082 TGATGAGCAGGTTCAGAGAGCTTGGCTACTCGGTGACTGACGCCAAGGGAATCTTCCA 1141
DB 723 TGATGAGCAGGTTCAGAGAGCTTGGCTACTCGGTGACTGACGCCAAGGGAATCTTCCA 782
OY 1142 GGTGTGATTTCAAGATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCAT 1201
DB 783 GGTGTGATTTCAAGATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCAT 842
OY 1202 GGACACTCTGGAGAAGAAAGACTTCTCTGAAG 1232
DB 843 GGACACTCTGGAGAAGAAAGACTTCTCTGAAG 873

RESULT 7
AAK94713
ID AAK94713 standard; cDNA; 2702 BP.
XX AAK94713;
AC AAK94713;
XX 06-NOV-2001 (first entry)
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3751.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW Homo sapiens.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
XX 05-SEP-2001.
PD 05-SEP-2001.
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogat T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI: 2001-524255/58.
DR P-PSDB: AAM93760.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3751; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 2702 BP; 608 A; 757 C; 699 G; 638 T; 0 other;

Query Match		68.9%;	Score 848.8;	DB 22;	Length 2702;
Best Local Similarity		99.8%;	Pred. No. 5.9e-206;		
Matches 850;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	381 ATGGCTGGGAGGGCCCCGGCGCCGACAGCCGGAGGACAGAGAGGAGCGGGTCTGAGAGCTT	440			
Db	1 ATGGCTGGGAGGGCCCCGGCGCCGACAGCCGGAGGACAGAGAGGAGCGGGTCTGAGAGCTT	60			
QY	441 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTTAGACATCAACATCCAGGAACCT	500			
Db	61 CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTTAGACATCAACATCCAGGAACCT	120			
QY	501 CGCTGGGACCAAAAGTACTTCTCTGGGGCAGAGCCCGGACATTTTCACTGTTACTGATCCT	560			
Db	121 CGCTGGGACCAAAAGTACTTCTCTGGGGCAGAGCCCGGACATTTTCACTGTTACTGATCCT	180			
QY	561 GGAATCTGCTGCTGTCCGGGGCAGCTGGAAGCTTCTCGGACATCGTGACAGAACTAC	620			
Db	181 GGAATCTGCTGCTGTCCGGGGCAGCTGGAAGCTTCTCGGACATCGTGACAGAACTAC	240			
QY	621 AGGCGCGCGCTGTGTGACCCGAGGATCACCGAGACACAGCTGTGAGGGCCAACTATGTG	680			
Db	241 AGGCGCGCGCTGTGTGACCCGAGGATCACCGAGACACAGCTGTGAGGGCCAACTATGTG	300			
QY	681 TATGACTCCGCTTCCATCCGACACAGGGGAGAGAGGTGTCTGATTTGGCCGATGTCA	740			
Db	301 TATGACTCCGCTTCCATCCGACACAGGGGAGAGAGGTGTCTGATTTGGCCGATGTCA	360			
QY	741 GCCCAGGTGCCCATGAACATGACATCACTGGCTGCATGCTCACAATTCTACAGGAAGACC	800			
Db	361 GCCCAGGTGCCCATGAACATGACATCACTGGCTGCATGCTCACAATTCTACAGGAAGACC	420			
QY	801 CCAACCGTGTGTCTGTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTACTACTCC	860			
Db	421 CCAACCGTGTGTCTGTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTACTACTCC	480			
QY	861 AACCGCAGTGTGTGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTAGTGGC	920			
Db	481 AACCGCAGTGTGTGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTAGTGGC	540			
QY	921 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCC	980			
Db	541 ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCC	600			
QY	981 TTGCTCGGCAGATTGTGCCCTTTTGACAGCAGTGGCAGCTGCCCAACTGCATCAACATCCC	1040			
Db	601 TTGCTCGGCAGATTGTGCCCTTTTGACAGCAGTGGCAGCTGCCCAACTGCATCAACATCCC	660			
QY	1041 CTGATGAGGACAGAGAGAGCTGACAGGTGGGCATCCCGGTGGCTGATGAGGACAGGTACAGG	1100			
Db	661 CTGATGAGGACAGAGAGAGCTGACAGGTGGGCATCCCGGTGGCTGATGAGGACAGGTACAGG	720			
QY	1101 CTTGGCTACTCGGTGACTGACGCCCAAGCAGGGAATCTTCCAGGTGATTTCAAGAATC	1160			
Db	721 CTTGGCTACTCGGTGACTGACGCCCAAGCAGGGAATCTTCCAGGTGATTTCAAGAATC	780			
QY	1161 TGCATGGCGATTCTTGCATGGCCATCCACCACACTGATCATGACACTCTGAGAGAAGAA	1220			
Db	781 TGCATGGCGATTCTTGCATGGCCATCCACCACACTGATCATGACACTCTGAGAGAAGAA	840			
QY	1221 GACTTCTCTGAAG 1232				
Db	841 GACTTCTCTGAAG 852				

RESULT 8
ABL39805
ID ABL39805 standard; cDNA; 984 BP.

XX ABL39805;
AC
XX
DT 10-MAY-2002 (first entry)
XX

DE	Human NS cDNA sequence SEQ ID NO:115.
XX	
KW	Human; cytosolic; osteopathic; gynaecological; neuroprotective;
KW	antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW	vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW	anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
KW	antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW	anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuroleptic;
KW	gastrointestinal; virucide; antilicer; cerebroprotective; nootropic;
KW	contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW	endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW	rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW	infertility; cardiovascular disease; coagulation disease; hypertension;
KW	ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW	diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW	gastric ulcer; Alzheimer's disease; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200206315-A2.
XX	
PD	24-JAN-2002.
XX	
PF	17-JUL-2001; 2001WO-IL00653.
XX	
PR	18-JUL-2000; 2000IL-0137345.
PR	15-DEC-2000; 2000IL-0140354.
XX	
PA	(COMP-) COMPUGEN LTD.
XX	
PI	Mintz L, Freilich S, Bernstein J;
XX	
DR	WPI; 2002-155037/20.
DR	P-PSDB; ABB06151.
XX	
PT	One hundred and twenty eight novel nucleic acid sequences, useful for
PT	treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX	
PS	Claim 1; Page 141-142; 290pp; English.
XX	
CC	ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC	encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC	(NS) can have cytosolic, osteopathic, gynaecological, neuroprotective,
CC	antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC	vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC	anorectic, muscular, anti-HIV, antinfertility, cardiovascular,
CC	anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC	immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antilicer,
CC	antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC	nootropic and contraceptive activities. The NS can be used in vaccines,
CC	gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC	antibodies from the present invention can be used for treating and
CC	diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC	diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC	cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC	glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC	disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC	disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC	depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC	Alzheimer's disease and as a contraceptive.
XX	
SO	Sequence 984 BP; 224 A; 278 C; 285 G; 196 T; 1 other;

Query Match 66.4%; Score 818.6; DB 24; Length 984;
Best Local Similarity 99.8%; Pred. No. 2.1e-198;
Matches 829; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY	402 CGACAGCGGAGGACAGAGGAAGCGGTTCTGAGAGCTTCAGAGAGCGATGAAAGCAAA	461
Db	1 CGACAGCGGAGGACAGAGGAAGCGGTTCTGAGAGCTTCAGAGAGCGATGAAAGCAAA	60
QY	462 ATGGTGAATTGCTTTAGACATCAACATCCAGGAACCTGCTGGACCAAAAGTACTTTC	521

Db 61 ATGGGTGAATGGCTTTAGACATCAACATCCAGGAACCTCGCTGGAGCCAAAGTACTTTC 120
QY 522 CTGGCAGAGCCCGGACATTTTTCATCTGTACTGTATCCTCGAAATCTGCTGCTGTCGGG 581
Db 121 CTGGGCAGAG-CCGGCAGCTTTTCACTGTACTGATCCTCGAAATCTGCTGCTGTCGGG 179
QY 582 GCACAGCTGGAAGCTTCTCGAACAATCGTGACAAGTACAGGGCGCGCTGTGACCCCA 641
Db 180 GCACAGCTGGAAGCTTCTCGAACAATCGTGACAAGTACAGGGCGCGCTGTGACCCCA 239
QY 642 GGGATCACCGAGGAGCAGCTGTGAGGGCCCAAGTATGTATGACTCCGCTTCATCCG 701
Db 240 GGGATCACCGAGGAGCAGCTGTGAGGGCCCAAGTATGTATGACTCCGCTTCATCCG 299
QY 702 GACACAGGGGAGAAGGTGCTCTGATTTGGCCGATGTCAGCCCGAGGTGCCCATGAACATG 761
Db 300 GACACAGGGGAGAAGGTGCTCTGATTTGGCCGATGTCAGCCCGAGGTGCCCATGAACATG 359
QY 762 ACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTGTTCTGGCAG 821
Db 360 ACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTGTTCTGGCAG 419
QY 822 TGGGTGAATCAGTCTTCAATGCCATTGTAACTACTCCAAACCGAGTGTGACACTCCC 881
Db 420 TGGGTGAATCAGTCTTCAATGCCATTGTAACTACTCCAAACCGAGTGTGACACTCCC 479
QY 882 ATCACTGTGAGGAGCAGCTGGGAGACGCTATGTAGTGCCACCACTGGAGCTGTGGCCACG 941
Db 480 ATCACTGTGAGGAGCAGCTGGGAGACGCTATGTAGTGCCACCACTGGAGCTGTGGCCACG 539
QY 942 GCCCTGGGACTCAATCCCTCACACAGCACTGCCCCCTTGGTCGGCAGATTGTGCCCC 1001
Db 540 GCCCTGGGACTCAATCCCTCACACAGCACTGCCCCCTTGGTCGGCAGATTGTGCCCC 599
QY 1002 TTTGCAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGCAGAGAGAGCTG 1061
Db 600 TTTGCAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGCAGAGAGAGCTG 659
QY 1062 CAGGTGGGCATCCCGGTGGCTGATGAGGAGGCTCAGAGGCTTGCTACTCGGTGACTGCA 1121
Db 660 CAGGTGGGCATCCCGGTGGCTGATGAGGAGGCTCAGAGGCTTGCTACTCGGTGACTGCA 719
QY 1122 GCCAAGCAGGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCGATTCTGCCATG 1181
Db 720 GCCAAGCAGGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCGATTCTGCCATG 779
QY 1182 GCCATCCACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCTCTGAAG 1232
Db 780 GCCATCCACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCTCTGAAG 830

RESULT 9
AAC75798
ID AAC75798 standard; cDNA; 843 BP.
AC AAC75798;
XX
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF1353 polynucleotide sequence SEQ ID NO:2705.
XX

KW Human; open reading frame; ORFX; detection; cytosstatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX Homo sapiens.
OS WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
PF 31-MAR-1999; 99US-0127607.
XX
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
PI
XX
XX MPI: 2000-602362/57.
DR P-PSDB; AAB41589.
DR
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
PS Claim 5; Page 1942; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytosstatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 843 BP; 183 A; 241 C; 252 G; 165 T; 2 other;
SQ

Query Match 65.4%; Score 805.8; DB 21; Length 843;
Best Local Similarity 97.4%; Pred. No. 3.5e-195;
Matches 819; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 362 ACCGCTGACGTCCCGCGTGATGGCTGGAGGGCCCGGGCGACACCGGAGGAGAGG 421
Db 3 ACCGCTGACGTCCCGCGTGATGGCTGGAGGGCCCGGGCGACACCGGAGGAGAGG 62
QY 422 AAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTGCCCTTAGA 481
Db 63 AAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTGCCCTTAGA 122
QY 482 CATCAACATCCAGGAACCTGCTGGAGCCAAAGTACTTTCTGGCAGAGCCCGGCACTT 541
Db 123 CATCAACATCCAGGAACCTGCTGGAGCCAAAGTACTTTCTGGCAGAGCCCGGCACTT 182
QY 542 TTTCACTGTACTGATCCTCGAAATCTGCTGCTCCGGGGCAGACGCTGAAGCTTCTCG 601
Db 183 TTTCACTGTACTGATCCTCGAAATCTGCTGCTCCGGGGCAGACGCTGAAGCTTCTCG 242

QY 602 GAACATCGTCAGAACTACAGGGCGCGGTGGTGACCCAGGGATCACCAGAGACGACT 661
DB 243 GAACATCGTCAGAACTACAGGGCGCGGTGGTGACCCAGGGATCACCAGAGACGACT 302
QY 662 GTGAGGGCCAAAGTATGTATGACTCCGCCCTTCCATCCGGACACAGGAGGAGAGGTGT 721
DB 303 GTGAGGGCCAAAGTATGTATGACTCCGCCCTTCCATCCGGACACAGGAGGAGAGGTGT 362
QY 722 CCTGATTGGCCGATGTCAGCCCAAGTGGCCCATGAACATGACCATCAGTGGCTGCATGCT 781
DB 363 CCTGATTGGCCGATGTCAGCCCAAGTGGCCCATGAACATGACCATCAGTGGCTGCATGCT 422
QY 782 CACATTTCTACAGGAAGACCCCAACCGTGGTCTCTGGCAGTGGGTGAATCAGTCTTCAA 841
DB 423 CACATTTCTACAGGAAGACCCCAACCGTGGTCTCTGGCAGTGGGTGAATCAGTCTTCAA 482
QY 842 TGGCATTTGTTAACTACTCCAAACCGCAGTGGTGACACTCCCATCTGTGAGGACAGTGG 901
DB 483 TGGCATTTGTTAACTACTCCAAACCGCAGTGGTGACACTCCCATCTGTGAGGACAGTGG 542
QY 902 GACAGCCTATGTAGTGGCAACCACTGGAGCTGTGGCCAGGCCCTGGACTCAAAATCCCT 961
DB 543 GACAGCCTATGTAGTGGCAACCACTGGAGCTGTGGCCAGGCCCTGGACTCAAAATCCCT 602
QY 962 CACCAAGCACCTGCCCCCTTGGTCGGCAGATTTGGCCCTTTGCAGCAGTGGCAGTGC 1021
DB 603 CACCAAGCACCTGCCCCCTTGGTCGGCAGATTCGTACCTTTGCAGCAGTGGCAGTGC 662
QY 1022 CAACCTGCATCAACATCCCCCTGATGAGGACAGAGAGAGAGTGCAGGTGGGCATCCC 1081
DB 663 CAACCTGCATCAACATCCCCCTGATGAGGACAGAGAGAGAGTGCAGGTGGGCATCCC 722
QY 1082 TGATGAGGCGAGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTCA 1141
DB 723 TGATGAGGCGAGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTCA 782
QY 1142 GGTGGTGAATTTCAAGAACTCTGATGGCGATTCTGCGCATGGCCATCCACACACTGATCAT 1201
DB 783 GGTGGTGAATTTCAAGAACTCTGATGGCGATTCTGCGCATGGCCATCCACACACTGATCAT 842
QY 1202 G 1202
DB 843 G 843

RESULT 10
AAK91763
ID AAK91763 standard; cDNA; 752 BP.
XX
AC AAK91763;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA 5'-end sequence, SEQ ID NO: 223.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 2; SEQ ID NO 223; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC sequence of the 5'-end of a cDNA provided in the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 752 BP; 173 A; 208 C; 214 G; 149 T; 8 other;

Query Match 53.5%; Score 659; DB 22; Length 752;
Best Local Similarity 96.5%; Pred. No. 8e-158;
Matches 713; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 381 ATGGCTGGAGGGCGCGCGCGACAGCGGAGGACAGAGGAAGCGGTCTGAGAGCT 440
DB 1 ATGGCTGGAGGGCGCGCGCGACAGCGGAGGACAGAGGAAGCGGTCTGAGAGCT 60
QY 441 CAGAGAGCGATGAAAGCAAAATGGGTGAATTGCCCTTGAACATCAACATCCAGAACT 500
DB 61 CAGAGAGCGATGAAAGCAAAATGGGTGAATTGCCCTTGAACATCAACATCCAGAACT 120
QY 501 CGCTGGGACCAAGTACTTCCCTGGGCGAGAGCCCGGACTTTTCACTGTACTGATCCT 560
DB 121 CGCTGGGACCAAGTACTTCCCTGGGCGAGAGCCCGGACTTTTCACTGTACTGATCCT 180
QY 561 CGAATCTGCTGCTGCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 620
DB 181 CGAATCTGCTGCTGCCGGGCGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTAC 240
QY 621 AGGGCCGGCGGTGTGACCCCA-GGGATCACCGAGGACCACTGTGAGGGCCCAAGTATGT 679
DB 241 AGGGCCGGCGGTGTGACCCCAAGGGGATCACCGAGGACCACTGTGAGGGCCCAAGTATGT 300
QY 680 GTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAGGTGCTCTGATTGGCCGATGTC 739
DB 301 GTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAGGTGCTCTGATTGGCCGATGTC 360
QY 740 AGCCGAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTACATTTACAGGAAGAC 799
DB 361 AGCCGAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTACATTTACAGGAAGAC 420
QY 800 CCCAACCCTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAACTACTC 859
DB 421 CCCAACCCTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAACTACTC 480
QY 860 CAACCGCAGTGTGACACTCCCATCACTGTGAGGACAGTGGGGACAGCCTATGTAGTGC 919
DB 481 CAACCGCAGTGTGACACTCCCATCACTGTGAGGACAGTGGGGACAGCCTATGTAGTGC 540
QY 920 CACCACTGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTACCAAGCACCTGCCCC 979
DB 541 CACCACTGAGCTGTGGCCACGGNCCTGGGACTCAAAATCCCTACCAAGCACCTGCCCC 600
QY 980 CTGGGTGGCAGATTTGTGCCCTTTGACAGCAGTGGCAGCTGCCAAC-TGCATCAACATCC 1038
DB 601 CTGGGTGGCAGATTTGTGCCCTTTGACAGCAGTGGCAGCTGCCAACCTTGCATCAACATCC 660
QY 1039 CCTGATGAGGACAGAGAGAGCTGC-AGGTGGCATCCCGGTGG--CTGATGAGGACAGTGC 1095

Db	661	CCCTGATGAGCANAANAACATGCAAGGTNGCATCCCAATGGCGTGATGAAGCAGGTC	720
QY	1096	AGAGCCTTGCTACTCGGT	1114
Db	721	AAAAGCTTGCTACTCCGT	739
<p>RESULT 11</p> <p>AAK93199</p> <p>ID AAK93199 standard; cDNA; 752 BP.</p> <p>XX AAK93199;</p> <p>XX</p> <p>DT 06-NOV-2001 (first entry)</p> <p>XX</p> <p>DE Human cDNA clone representative sequence, SEQ ID NO: 1659.</p> <p>XX</p> <p>KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.</p> <p>XX</p> <p>OS Homo sapiens.</p> <p>XX</p> <p>PN EP1130094-A2.</p> <p>XX</p> <p>PD 05-SEP-2001.</p> <p>XX</p> <p>PF 07-JUL-2000; 2000EP-0114089.</p> <p>XX</p> <p>PR 08-JUL-1999; 99JP-0194486.</p> <p>PR 11-JAN-2000; 2000JP-0118774.</p> <p>PR 02-MAY-2000; 2000JP-0183765.</p> <p>XX</p> <p>PA (HELI-) HELIX RES INST.</p> <p>XX</p> <p>PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;</p> <p>PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;</p> <p>XX</p> <p>DR WPI; 2001-524255/58.</p> <p>XX</p> <p>PT 830 Primers useful for synthesizing full length cDNA clones and their</p> <p>PT use in genetic manipulation -</p> <p>XX</p> <p>PS Example 11; SEQ ID NO 1659; 1380pp + sequence listing; English.</p> <p>XX</p> <p>CC The invention relates to primers for synthesizing full length cDNA</p> <p>CC clones. 830 cDNA molecules encoding a human protein have been</p> <p>CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA</p> <p>CC molecules have been determined. Primers for synthesizing the full length</p> <p>CC cDNA are useful for clarifying the function of the protein encoded by</p> <p>CC the cDNA. The full length clones were obtained by construction of full</p> <p>CC length enriched cDNA libraries that were synthesised by the oligo-capping</p> <p>CC method. The primers enable the production of the full length cDNA easily</p> <p>CC without any special methods. The present sequence was used as the</p> <p>CC representative sequence from a human clone which was used in</p> <p>CC homology searches to identify the clone.</p> <p>CC Note: The sequence data for this patent did not form part of the printed</p> <p>CC specification, but was obtained in CD-ROM format directly from EPO.</p> <p>CC</p> <p>XX</p> <p>XX</p> <p>SEQ Sequence 752 BP; 173 A; 208 C; 214 G; 149 T; 8 other;</p>			
QY		Query Match	53.5%; Score 659; DB 22; Length 752;
		Best Local Similarity	96.5%; Pred. No. 8e-158;
		Matches 713; Conservative	0; Mismatches 21; Indels 5; Gaps 4
Db	1	ATGGCTGGAGGGCCCGCGCGACACGCGAGGACAGAGAGGAAGCGGCTTGAGAGCTT	440
QY	381	ATGGCTGGAGGGCCCGCGCGACACGCGAGGACAGAGAGGAAGCGGCTTGAGAGCTT	440
Db	1	ATGGCTGGAGGGCCCGCGCGACACGCGAGGACAGAGAGGAAGCGGCTTGAGAGCTT	60
QY	441	CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCT	500
Db	61	CAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCT	120
QY	501	CGCTGGAGCAACAGTACTTTCCTGGGCGAGAGAGCCCGGACATTTCACCTGTACTGATCCT	560

Db	121	CGCTGGGACCAAGTACTTTCTCGGGCAGAGCCCGGCACTTTTCACTGTTACTGATCCT	180
QY	561	CGAATCTGCTGCTGTCGCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTAC	620
Db	181	CGAATCTGCTGCTGTCGCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTAC	240
QY	621	AGGGCCGGCGTGGTGAACCCA - GGGATCACCCGAGGACAGCTGTGGAGGGCCAGTATGT	679
Db	241	AGGGCCGGCGTGGTGAACCCAAGGGATCACCCGAGGACAGCTGTGGAGGGCCAGTATGT	300
QY	680	GATGACTCCGCTTCATCCGGACACAGGGGAGAGGTGCTCTGATTGGCCGATGTC	739
Db	301	GATGACTCCGCTTCATCCGGACACAGGGGAGAGGTGCTCTGATTGGCCGATGTC	360
QY	740	AGCCCAAGGTGCCCATGACATGACCATCACTGGCTGCATGCTCACATTCTACAGAAAGAC	799
Db	361	AGCCCAAGGTGCCCATGACATGACCATCACTGGCTGCATGCTCACATTCTACAGAAAGAC	420
QY	800	CCCAACCGTGTGTCTCGGACAGTGGGTGAATCACTCCCTCAATGCCATTGTTAACTACTC	859
Db	421	CCCAACCGTGTGTCTCGGACAGTGGGTGAATCACTCCCTCAATGCCATTGTTAACTACTC	480
QY	860	CAACCGCAGTGTGTGACACTCCCATCACTGTGAGGCAAGCTGGGACAGCCCTATGTGAGTGC	919
Db	481	CAACCGCAGTGTGTGACACTCCCATCACTGTGAGGCAAGCTGGGACAGCCCTATGTGAGTGC	540
QY	920	CACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCC	979
Db	541	CACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCC	600
QY	980	CTTGGTCGGCAGATTGTGCCCTTTGCAGCAGTGCAGCTGCCAAC - TGCATCAACATCC	1038
Db	601	CTTGGTCNGCAGATTGTGCCCTTTGCAGCAGTGCAGCTGCCAACCTGCACACTTGCATCAACATCC	660
QY	1039	CCCTGATGAGGCAGAGAGAGCTGC - AGGTGGGCAATCCCGGTGG - CTGATGAGGCAGGTC	1095
Db	661	CCCTGATGAGGCAGAGAGAGCTGCAGAGGTNGGCATCCCAATGGGCTGATGAAGGCAGGTC	720
QY	1096	AGAGGCTTGGCTACTCGGT	1114
Db	721	AAAAGCTTGGCTACTCCGT	739
RESULT 12			
AAS84362			
ID	AAS84362	standard; cDNA; 2215 BP.	
XX	AAS84362;		
AC			
XX			
DT	13-FEB-2002	(first entry)	
XX			
DE		DNA encoding novel human diagnostic protein #20166.	
XX			
KW		Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX		food supplement; medical imaging; diagnostic; genetic disorder; ss.	
OS		Homo sapiens.	
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001; 2001WO-US08631.		
XX			
PR	31-MAR-2000; 2000US-0540217.		
XX			
PR	23-AUG-2000; 2000US-0649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Dymanac RT, Liu C, Tang YT;		
XX			
DR	WPI; 2001-639362/73.		

DR P-PSDB; ABG20175.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID NO 20166; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2215 BP; 491 A; 649 C; 582 G; 492 T; 1 other;

Query Match 51.3%; Score 632.2; DB 23; Length 2215;
Best Local Similarity 90.9%; Pred. No. 7.5e-151;
Matches 766; Conservative 0; Mismatches 8; Indels 69; Gaps 6;

QY 459 AAAATGGGTGAATTGCTTTAGACATCAATCCAGGAACCTCGCTGGGACCAAGTACT 518
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
208 AAAATGGGTGAATTGCTTTAGACATCAATCCAGGAACCTCGCTGGGACCAAGTACT 267
QY 519 TTCTTGGGCAAGC-CCGGCACCTTTTCACTGTACTGATCCTCGAAATCTGCTGTC 577
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
268 TTCTTGGGCAAGC-CCGGCACCTTTTCACTGTACTGATCCTCGAAATCTGCTGTC 327
QY 578 CGGGCACAGCTGGAAGCTTCTCGAATCGTGAGAACTACAGGCGCGGTGTGAC 637
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
328 CGGGCACAGCTGGAAGCTTCTCGAATCGTGAGAACTACAGGCGCGGTGTGAC 387
QY 638 CCCAGGATCACCGAGACAGCTGTGGAGGGCCAAAGTATGTATGACTCCGCTTCCA 697
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
388 CCCAGGATCACCGAGACAGCTGTGGAGGGCCAAAGTATGTATGACTCCGCTTCCA 447
QY 698 TCCGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTCAAGCCCAAGTGCCCATGAA 757
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
448 TCCGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTCAAGCCCAAGTGCCCATGAA 507
QY 758 CATGACCATCACTGGCTGCATGCTCAATTTA----- 790
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
508 CATGACCATCACTGGCTGCATGCTCAATTTA----- 567
QY 791 -----CAGGAAGACCCCAACCGTGTGTT 814
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
568 CTGTAGACGGGGCAGAGTGAAGTGTCTTGTTCCTCAGGAAGACCCCAACCGTGTGTT 627
QY 815 CTGGCAGTGGGTGAATCAGTCTCTCAATGCCATTTGTTAACTACTCCAAACCGCA-GTGGTG 873
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
628 CTGGCAGTGGGTGAATCAGTCTCTCAATGCCATTTGTTAACTACTCCAAACCGCAAGTGGTG 687
QY 874 ACACCT--CCCATCACTGTGAGGACAGCTGGGGA-CAGCCTATGTGAGTGCCA-CCACTGGA 929
Db ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
688 ACACCTCCCATCACTTGTGAGGACAGCTGGGGAACCAACCTATGTGAGTGCCANCCACTGGA 747

QY 930 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGTCGGC 989
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
748 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGTCGGC 807
QY 990 AGATTGTGCTTGTGACAGAGTGGAGCTGCCAATGATCAATCCCTGATGAGG 1049
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
808 AGATTGTGCTTGTGACAGAGTGGAGCTGCCAATGATCAATCCCTGATGAGG 867
QY 1050 CAGAGAGAGCTGACAGTGGGATCCCGGTGCTGATGAGGACAGTCCAGAGCTTGCTAC 1109
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
868 CAGAGAGAGCTGACAGTGGGATCCCGGTGCTGATGAGGACAGTCCAGAGCTTGCTAC 927
QY 1110 TCGGTGACTGCAGCCAGCAGGAATCTCCAGGTGTGATTTCAAGAATCTGCATGGCG 1169
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
928 TCGGTGACTGCAGCCAGCAGGAATCTCCAGGTGTGATTTCAAGAATCTGCATGGCG 987
QY 1170 ATTCCTGACATGGCCATCCCACTGATCATGACACTCTGAGAGAAAGACTTCCTG 1229
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
988 ATTCCTGACATGGCCATCCCACTGATCATGACACTCTGAGAGAAAGACTTCCTG 1047
QY 1230 AAG 1232
Db |||
1048 AAG 1050

RESULT 13
AAI60909
ID AAI60909 standard; cDNA; 723 BP.

XX AAI60909;

DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 4898.

KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

OS Homo sapiens.
XX
PN WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41753.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS Claim 1; SEQ ID NO 4898; 10078bp; English.
XX

QY	616	ACTACAGGGGCGCGGTGGTGATCCCAAGGGATCACCAGGACAAGCTGTGGAGGGCCAAGT	675
Db	282	ATTACAGGCAGGGAATTGTTCTCTCTGCTCTTACAGAATAATTTGTGGAGAGCAAAAGT	341
QY	676	ATGTGTATGACTCCGCCCTTCATCCGGACACAGGGGAGAAAGTGTCTGATTTGCCGCA	735
Db	342	ACATCTATGATTACAGCTTTTCATCTCTGACACTGGTAGAAGATGATTTTGATAGGAAGAA	401
QY	736	TGTCAGCCCCAGGTGCCCATGACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGA	795
Db	402	TGTCAGCCCAGGTTCCCATGACATGACCATCACAGGTTGTATGATGACGTTTTACAGGA	461
QY	796	AGACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCCATTTGTTAACT	855
Db	462	CTACGCGCGGTGTGCTGTGTTCTGGCAGTGGATTAACCAAGTCTTCAATGCCGCTGTCATAT	521
QY	856	ACTCCAACCGCAGTGTGTGACACTCCCATCACTGTGAGGCAAGCTGGGGACAGCCTATGTGA	915
Db	522	ACACCAACAGAAGTGGAGACGCAACCCCTCACTGTCATGAGCTTGGGAACAGCTTACGTTT	581
QY	916	GTCGACACCACGTGAGCTGTGGCCACAGGCCCTGGGACTCAAATCCCTCACCAAGCACCTGC	975
Db	582	CTGCAACAACACTGTGTGCCGTAGCAACAGCTCTAGAGCTCAATGATTTGNACCAAGCATGTCT	641
QY	976	CCCCCTTGGTCGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCTGCAACTGCATCAACA	1035
Db	642	CACCACTGATAGGACGTTTTGTTCCCTTTGCTGCCGTAGCTGCTTAATTGCAATTAATA	701
QY	1036	TCCCCCTGATGAGGCAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAGGTC	1095
Db	702	TTCCATTAATGAGGCAAGGGAACCTCAAAAGTTGGCATTTCCGTCACGGATGAGAATGGGA	761
QY	1096	AGAGGCTTGGCTACTCGGTGACTGCAAGCCAAAGCAGGGAATCTTCCAGGTGTGATTCAA	1155
Db	762	ACCGCTTGGGGGAGTCCGCGCAACGCTGCCGAACAAGCCATCACGCCAAGTTGTGCTGTCCA	821
QY	1156	GAATCTGCATGGCCGATTCTCTGCCATGGCCATCCCAACCACTGATCATGGACACTCTGGAGA	1215
Db	822	GGATTCTCATGGCAGCCCTGGCATGGCCATCCCTCCATTCATTATGAACACACTTTGGAAA	881
QY	1216	AGAAAGACTTCCTGAAG 1232	
Db	882	AGAAAGCCTTTTGAAG 898	

RESULT 15	
AAI59569	
ID	AAI59569 standard; cDNA; 1076 BP.
XX	
AC	AAI59569; 1
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 1772.
XX	
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR P-PSDB; AAM40413.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1772; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAW38642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ sequence 1076 BP; 285 A; 258 C; 269 G; 264 T; 0 other;

Query Match	32.38;	Score 398.2;	DB 22;	Length 1076;
-------------	--------	--------------	--------	--------------

Best Local Similarity 69.18; Pred. No. 2.2e-91;

Matches 544; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

Qy 446 AGCGATGGAAGCAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGGAACCTCGCTG 509
||| | | | | | | | | | | | | | | | |
Db 48 AGCGTCCGGGACCATTGTCTGGAGAACTACCAACAACATTAAACATCAAGGAACCTCGATG 107

QY 506 GGACCAAGTACTTCCTGGGCGAGCCCGGCACCTTTTCTACTGTTA CTGATCCTCGAA 5655

Dd 108 GGATCAAGCACCTTTCATTGGACGAGCCAATCATTTCTTCACTGTTA ACTGACCCCGA 1677

Qy 566 TCTGCTGCTGTCCGGGGCAGACGTTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGCC 6295
| | | | | | | | | | ~ | | | | | | | |
Db 168 CATTTCTGTTAACCAACGACAACACTCGAGAGTGCGAGAAAAATAGTACATGATTACAGGCA 2277

Dy 626 CGCGCTGCTGAACCCAGGGATCACCGAGCAACCAGCTGTGGAGGGCCAAGTATGTGATGA 6895
|| || | || | | || | | | ||||| || ||||| | |||||
Db 228 AGGAATTGTTCTCCTCGTGCTTACAGAAATGAATTGTGGAGAGCAAAGTACATCTATGA 2877

Dy 686 CTCCGCCCTTCATCCGGACACAGGGAAGGTGGTTCCTGATTGGCGGCATGTCAGGCCCA 745
 || || | | ||||| ||||| | | ||||| || | ||||| |||||

Db 288 TTCAGCTTTTCATCCTGACACTGGTGAGAAGATGATTTTGATAGGAAGAAATGTCAGGCCA 347

Oy 746 GGTGCCATGAACATGACCATCACTGGCTGCATGCTCACATTCTACAGGAAGACCCCAAC 805
||| |||||||||||||||| | | | | | | | | | | | |
db 348 GGTTCCCATGAACATGACCATCAAGGTTGTATGATGACGTTTTACAGGACTACGCCGGC 407

Dy 806 CGTGTCTCTGGCAGTGGGTAATCAGTCCTTCAATGCATTGTTAACTACTCCAACCG 865

Dd 408 TGTGCTGTTCTGGCAGTGGATTAACCACTCTTCAATGCCCTGCTCAATTACACCACAACAG 467

Qy 866 CAGTGGTGCACACTCCCATCACTGTGAGGGCAGCTGGGGACAGCCTATGTGAGTGCACACC 925
||||| ||| | ||| |||||| | || |||| |||||| || | ||| |||

Db	468	AA GTGAGACGCA CCCC TCAC TGTCAATGAGTGGGAA CAGCTTACGTTTCGACAAC	527
QY	926	TGGAGCTGTGGCCACG GCGCC TGGACTCAAA TCCCTCACC AAGCACC TGCCGCCCTTGGT	985
Db	528	TG GTGCCGTAGCAACAGCTCTAGGACTCAATGCAT TGACCAAGCATGTCTACCACTGAT	587
QY	986	CGGAGATTGTGCCCC TTTGCAGCAGTGGCAGCTGCCA ACTGCATCAACATCCCCCTGAT	1045
Db	588	AGGACGTTTGTTCCTTCCCTT TGTGTCGCCGTAGCTGCTGCTAATTGCATTAA TTTCCATTAA T	647
QY	1046	GAGCAGAGAGAGACTG CAGGTGGGCATCCCGGTGCGCTGATGAGGCAGGTCAGAGCCTTGG	1105
Db	648	GAGGCAAAAGGGA CTCAAGTTGGCATTTCCCGTCACGGATGAGAA TGGGAACCGCCTTGGG	707
QY	1106	CTACTCGGTGACTGCAGCC AAGCAGGGAATCTTCCAGGTGGTGATTTCAAGATCTGCAT	1165
Db	708	GGAGTGC GCGCAACGCTGCGAAACAAGCATCAGCAAGTTGTCGTGCAGGATTTCTCAT	767
QY	1166	GGCGATTCTGCCATGGCCATCC CACCACATGATCATGGACACTCTGGAGAGAAAGACTT	1225
Db	768	GGCAGCCCGCTGGCATGGCCATCCCTCCATTCATTATGAA CACTTTGGAAAGAAAGCCCTT	827
QY	1226	CCTGAAG 1232	
Db	828	TTTGAAG 834	

Search completed: October 16, 2003, 14:00:18
Job time : 372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 13:54:01 : Search time 360 Seconds
(without alignments)
8987.194 Million cell updates/sec

Title: US-09-990-415A-1
Perfect score: 1232
Sequence: 1 cccttagcgccagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues
Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1232	100.0	1232	10 US-09-990-415A-1	Sequence 1, Appl1
2	944	76.6	1729	14 US-10-037-270-1016	Sequence 1016, Ap
3	880	71.4	1061	10 US-09-990-415A-3	Sequence 3, Appl1
4	869.4	70.6	1375	9 US-09-867-550-1343	Sequence 1343, Ap
5	399.4	32.4	2269	10 US-09-990-415A-7	Sequence 7, Appl1
6	398.2	32.3	2978	14 US-10-198-846-9767	Sequence 9767, Ap
7	396.8	32.2	1134	14 US-10-014-338-1	Sequence 1, Appl1
8	335.4	27.2	1567	10 US-09-990-415A-5	Sequence 5, Appl1
9	332.2	27.0	2376	13 US-10-098-841-64	Sequence 64, Appl1
10	330.6	26.8	2458	9 US-09-822-849A-490	Sequence 490, App
11	315	25.6	984	14 US-10-014-338-3	Sequence 3, Appl1
12	272.8	22.1	670	14 US-10-206-901B-41	Sequence 41, Appl1
13	212.4	17.2	489	11 US-09-918-995-23576	Sequence 23576, A
14	197.2	16.0	293	10 US-09-783-590-2110	Sequence 2110, Ap
15	166.2	13.5	3342	12 US-09-814-353-20153	Sequence 20153, A
16	131.2	10.6	522	11 US-09-918-995-5675	Sequence 5675, Ap

17	122	9.9	648	14 US-10-198-846-7595	Sequence 7595, Ap
18	118.2	9.6	496	14 US-10-106-698-2663	Sequence 2663, Ap
19	89	7.2	533	11 US-09-918-995-26900	Sequence 26900, A
20	73.4	6.0	238	11 US-09-991-936-1857	Sequence 1857, Ap
21	46.4	3.8	248	10 US-09-728-445-404	Sequence 404, App
22	41.6	3.4	2303	10 US-09-880-107-3849	Sequence 3849, Ap
23	41.6	3.4	2303	11 US-09-922-146-3	Sequence 3, Appl1
24	40.6	3.3	135638	12 US-10-314-657-1	Sequence 1, Appl1
25	40	3.2	2631	14 US-10-156-761-6208	Sequence 6208, Ap
26	40	3.2	9025608	14 US-10-156-761-1	Sequence 1, Appl1
27	39.8	3.2	5025	10 US-09-960-253-176	Sequence 176, App
28	39.8	3.2	42999	10 US-09-799-462A-17	Sequence 17, Appl
29	39.8	3.2	42999	11 US-09-836-911A-17	Sequence 17, Appl
30	39.8	3.2	42999	12 US-09-738-630-73	Sequence 73, Appl
31	39.8	3.2	42999	13 US-10-125-767-17	Sequence 17, Appl
32	39.8	3.2	42999	14 US-10-151-081-17	Sequence 17, Appl
33	39.8	3.2	42999	14 US-10-287-313-17	Sequence 17, Appl
34	39.8	3.2	42999	14 US-10-219-694-17	Sequence 17, Appl
35	39.4	3.2	390	14 US-10-156-761-44	Sequence 44, Appl
36	39.4	3.2	9025608	14 US-10-156-761-1	Sequence 1, Appl1
37	38.8	3.1	2561	10 US-09-976-740-48	Sequence 48, Appl
38	38.8	3.1	2561	13 US-10-023-529-48	Sequence 48, Appl
39	38.8	3.1	2561	13 US-10-023-523-48	Sequence 48, Appl
40	38.4	3.1	6453	12 US-09-960-706-618	Sequence 618, App
41	38	3.1	2000	10 US-09-887-576-874	Sequence 874, App
42	37.6	3.1	1874	10 US-09-323-998D-33	Sequence 33, Appl
43	37.4	3.0	838	13 US-10-027-632-163967	Sequence 163967, Sequence 163968, Sequence 5724, Ap
44	37.4	3.0	838	13 US-10-027-632-163968	
45	37.4	3.0	1683	14 US-10-156-761-5724	

ALIGNMENTS

RESULT 1
US-09-990-415A-1
; Sequence 1, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster 1
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (450)..(1232)
US-09-990-415A-1

Query Match	100.0%;	Score 1232;	DB 10;	Length 1232;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1232;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CCCTTAGGCGCCAGGACAGCCGAGCGTTACTGTCCCGGCGAGCGAGTCTTTACCC	60	
DB	1	CCCTTAGGCGCCAGGACAGCCGAGCGTTACTGTCCCGGCGAGCGAGTCTTTACCC	60	
QY	61	ACCCAGTCTGTCTGACGCGCCCTAGCTCATTCGCAAAATTAGGCTTGCTGGCT	120	
DB	61	ACCCAGTCTGTCTGACGCGCCCTAGCTCATTCGCAAAATTAGGCTTGCTGGCT	120	
QY	121	TGTTCCCTCCGCTCGAACACCTCTCTGTGAGCCGAGCCAGCTACCGGGCTCTGG	180	
DB	121	TGTTCCCTCCGCTCGAACACCTCTCTGTGAGCCGAGCCAGCTACCGGGCTCTGG	180	
QY	181	AATTGCCACCCCTCCCTGGGACACCTTGAGCCCTCCGTGAGGAGCTCACGGGCGAG	240	
DB	181	AATTGCCACCCCTCCCTGGGACACCTTGAGCCCTCCGTGAGGAGCTCACGGGCGAG	240	

```
Db 181 AATTGCCACCCCTCCCTGGGACACCTTGAAGCCCTCCGTGGAGGAGCTCACGGGGCAGAG 240
QY 241 CGGAGCTGAGCCTGAGTTTGTGAGAGCGCTGCTGTGTGTGGCTGGCTTCTGCAAT 300
Db 241 CGGAGCTGAGCCTGAGTTTGTGAGAGCGCTGCTGTGTGTGGCTGGCTTCTGCAAT 300
QY 301 CCCCGTCCACCGGGTGGCGCGCGGGAAGCTCCTGCCCCCTGCTGTGTGGCTG 360
Db 301 CCCCGTCCACCGGGTGGCGCGCGGGAAGCTCCTGCCCCCTGCTGTGTGGCTG 360
QY 361 CACGCTGAGCTCCCGCTGATGGCTGGAGGGCCCGCGCGCACAGCGGAGCAGAGAG 420
Db 361 CACGCTGAGCTCCCGCTGATGGCTGGAGGGCCCGCGCGCACAGCGGAGCAGAGAG 420
QY 421 GAAGCGGCTTCTGAGAGCTTCAAGAGCGGATGGAAGCAAAATGGGTGAATTGCCCTTAG 480
Db 421 GAAGCGGCTTCTGAGAGCTTCAAGAGCGGATGGAAGCAAAATGGGTGAATTGCCCTTAG 480
QY 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCTGGGCGAGAGCCCGGCACT 540
Db 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTTCTGGGCGAGAGCCCGGCACT 540
QY 541 TTTTCACTGTACTGATCCTGAAATCTGCTGCTGCTCCGGGCGACAGCTGGAAGCTTCTC 600
Db 541 TTTTCACTGTACTGATCCTGAAATCTGCTGCTGCTCCGGGCGACAGCTGGAAGCTTCTC 600
QY 601 GGAACATCGTGCAGAACTACAGGGCGCGCGCTGTGACCCCGAGGATCACCGAGGACCAAGC 660
Db 601 GGAACATCGTGCAGAACTACAGGGCGCGCGCTGTGACCCCGAGGATCACCGAGGACCAAGC 660
QY 661 TGTGGAGGGCCCAAGTATGTGTATGACTCCGCCCTTCATCCGGACACAGGGGAGAAGGTGG 720
Db 661 TGTGGAGGGCCCAAGTATGTGTATGACTCCGCCCTTCATCCGGACACAGGGGAGAAGGTGG 720
QY 721 TCCGTGATGGCCGATGTACAGCCAGGTGCCCATGAACATGACATCACTGGCTGCATGC 780
Db 721 TCCGTGATGGCCGATGTACAGCCAGGTGCCCATGAACATGACATCACTGGCTGCATGC 780
QY 781 TCACATTTCTACAGGAAGACCCCAACCGTGTGTCTGTGGCAGTGGGTGAATCAGTCCCTCA 840
Db 781 TCACATTTCTACAGGAAGACCCCAACCGTGTGTCTGTGGCAGTGGGTGAATCAGTCCCTCA 840
QY 841 ATGCCATTGTTAACTACTCTCAACCGCAGTGTGACACTCCATCACTGTGAGGCAGCTGG 900
Db 841 ATGCCATTGTTAACTACTCTCAACCGCAGTGTGACACTCCATCACTGTGAGGCAGCTGG 900
QY 901 GGACAGCCTATGTAGTGCACCACTGAGCTGTGGCCACGCGCTGGGACTCAAAATCCC 960
Db 901 GGACAGCCTATGTAGTGCACCACTGAGCTGTGGCCACGCGCTGGGACTCAAAATCCC 960
QY 961 TCACCAAGCACCTGCCCCCTTGGTGGCAGATTTGTGCCCTTTCAGCAGTGGCAGCTG 1020
Db 961 TCACCAAGCACCTGCCCCCTTGGTGGCAGATTTGTGCCCTTTCAGCAGTGGCAGCTG 1020
QY 1021 CCAACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGAGGTGGGCATCCCGGTGG 1080
Db 1021 CCAACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGAGGTGGGCATCCCGGTGG 1080
QY 1081 CTGATGAGGCGAGGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTTCC 1140
Db 1081 CTGATGAGGCGAGGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCCAAGCAGGGAATCTTCC 1140
QY 1141 AGGTGGTGAATTCAGAAATCTGCGATGGCGATTCCTGCGCATGCCATCCACCACTGATCA 1200
Db 1141 AGGTGGTGAATTCAGAAATCTGCGATGGCGATTCCTGCGCATGCCATCCACCACTGATCA 1200
QY 1201 TGGACACTCTGGAGAGAAAGACTTCTGAAG 1232
Db 1201 TGGACACTCTGGAGAGAAAGACTTCTGAAG 1232
```

RESULT 2
US-10-037-270-1016

```
Sequence 1016, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037, 270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (247)..(1371)
US-10-037-270-1016

Query Match 76.6%; Score 944; DB 14; Length 1729;
Best Local Similarity 100.0%; Pred. No. 1.2e-275;
Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GGTTCGCCAATCCCGGTGCCCCACCGGGTGGCGCGCGGGAAGCTCCTGCCCTCCCT 348
Db 86 GGTTCGCCAATCCCGGTGCCCCACCGGGTGGCGCGCGGGAAGCTCCTGCCCTCCCT 145
QY 349 GCTGTGGCGGTACACGCGGTGACGTCGCCGCTGATGGCTGGAGGGCCCGGGCGACAGC 408
Db 146 GCTGTGGCGGTACACGCGGTGACGTCGCCGCTGATGGCTGGAGGGCCCGGGCGACAGC 205
QY 409 GGAGGCGAGAGGAAGCGGCTTCTGAGAGCTTCAAGAGAGCGATGGAAGCAAAATGGGTG 468
Db 206 GGAGGCGAGAGGAAGCGGCTTCTGAGAGCTTCAAGAGAGCGATGGAAGCAAAATGGGTG 265
QY 469 AATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTCTGGGCA 528
Db 266 AATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGGACCAAAAGTACTTCTGGGCA 325
QY 529 GAGCCCGGCACTTTTTCATGTTACTGATCCTCGAAATCTGCTGTCCGGGGCACAGC 588
Db 326 GAGCCCGGCACTTTTTCATGTTACTGATCCTCGAAATCTGCTGTCCGGGGCACAGC 385
QY 589 TGGAGCTTCTCGAAGATCGTCAGAACTACAGGGCGCGGTGGTGAACCCAGGATCA 648
Db 386 TGGAGCTTCTCGAAGATCGTCAGAACTACAGGGCGCGGTGGTGAACCCAGGATCA 445
QY 649 CCGAGGACCAAGCTGTGAGGGGCAAGTATGTATGATACCTCGCCTTCCATCCGACACAG 708
Db 446 CCGAGGACCAAGCTGTGAGGGGCAAGTATGTATGATACCTCGCCTTCCATCCGACACAG 505
QY 709 GGGAGAAAGTGTCTCTGATTTGCCCGCATGTACGCCAGGTGCCCATGAACATGACATCA 768
```


|||||
Db 506 GGGAGAAAGGTGTCCTGATGGCCGCATGTCAGCCAGTGGCCCATGAACATGACCATCA 565
QY 769 CTGGCTGCATGCTCACATCTTACAGGAAGACCCCAACCGTGGTGTCTGAGTGGGTGA 828
Db 566 CTGGCTGCATGCTCACATCTTACAGGAAGACCCCAACCGTGGTGTCTGAGTGGGTGA 625
QY 829 ATCAGTCCCTTCAATGCCATGTGTAACACTCCAACCCGAGTGGTGACACTCCCATCACTG 888
Db 626 ATCAGTCCCTTCAATGCCATGTGTAACACTCCAACCCGAGTGGTGACACTCCCATCACTG 685
QY 889 TGAGGCAGCTGGGGACAGCCTATGTGAGTGCCACCACACTGGAAGCTGTGGCCACGGCCCTGG 948
Db 686 TGAGGCAGCTGGGGACAGCCTATGTGAGTGCCACCACACTGGAAGCTGTGGCCACGGCCCTGG 745
QY 949 GACTCAATCCCTCACCAAGCAACCTGCCCCCTTGGTGGCAGATTGTGCCCTTTCAG 1008
Db 746 GACTCAATCCCTCACCAAGCAACCTGCCCCCTTGGTGGCAGATTGTGCCCTTTCAG 805
QY 1009 CAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGCAGGTGG 1068
Db 806 CAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGCAGGTGG 865
QY 1069 GCATCCCGGTGCTGATGAGGCGAGGTCAGAGGCTTGGCTACTCGGTGACTGCAGCCAGC 1128
Db 866 GCATCCCGGTGCTGATGAGGCGAGGTCAGAGGCTTGGCTACTCGGTGACTGCAGCCAGC 925
QY 1129 AGGGAATCTTCCAGGTGTGATTTCAAGAATCTGCATGGCGATTCCTGCCATGGCCATCC 1188
Db 926 AGGGAATCTTCCAGGTGTGATTTCAAGAATCTGCATGGCGATTCCTGCCATGGCCATCC 985
QY 1189 CACCACTGATCATGAGCACTCTGGAGAAGAAAGACTTCCTGAAG 1232
Db 986 CACCACTGATCATGAGCACTCTGGAGAAGAAAGACTTCCTGAAG 1029

RESULT 3
US-09-990-415a-3
; Sequence 3, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (450)..(680)
; US-09-990-415a-3

Query Match 71.4%; Score 880; DB 10; Length 1061;
Best Local Similarity 86.1%; Pred. No. 2.6e-256;
Matches 1061; Conservative 0; Mismatches 0; Indels 171; Gaps 1;
QY 1 CCCTTAGGCGCCAGGAGACCGGAGCGTTACCTGTCCCGGGCAGCGAGTCTTTACCC 60
Db 1 CCCTTAGGCGCCAGGAGACCGGAGCGTTACCTGTCCCGGGCAGCGAGTCTTTACCC 60
QY 61 ACCCCAGTCTGTTCTGAGCGCCCTAGCTCATTCGCGAAATTAGGCTTGGGTGGCT 120
Db 61 ACCCCAGTCTGTTCTGAGCGCCCTAGCTCATTCGCGAAATTAGGCTTGGGTGGCT 120
QY 121 TGTTCCTCCGCTCGAACCACCTCTCTGTGAGCCGACCCAGCTACCGGGGCTCCTGG 180
Db 121 TGTTCCTCCGCTCGAACCACCTCTCTGTGAGCCGACCCAGCTACCGGGGCTCCTGG 180

QY 181 AATTGCCACCCCTCCCTGGGCAACCCCTTGAGGCCCTCCGTGAGGGACGTCACGGGGCAGAG 240
Db 181 AATTGCCACCCCTCCCTGGGCAACCCCTTGAGGCCCTCCGTGAGGGACGTCACGGGGCAGAG 240
QY 241 CGGAGCTGAGCCTGAGTTGGTGCAGGGCGGTGCTGTGTGGTGGCTTCTGCCAAT 300
Db 241 CGGAGCTGAGCCTGAGTTGGTGCAGGGCGGTGCTGTGTGGTGGCTTCTGCCAAT 300
QY 301 CCCCGTCCCAACCGGCTGGGCGGCGGGAAGCTCCTGCCCTCCCTGCTGTGGCGGT 360
Db 301 CCCCGTCCCAACCGGCTGGGCGGCGGGAAGCTCCTGCCCTCCCTGCTGTGGCGGT 360
QY 361 CAGCGGTGACGTCCCGGTGATGGCTGGGAGGGCCCGGGCGGACAGCGGAGCAGAGAG 420
Db 361 CAGCGGTGACGTCCCGGTGATGGCTGGGAGGGCCCGGGCGGACAGCGGAGCAGAGAG 420
QY 421 GAAGCGGCTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAG 480
Db 421 GAAGCGGCTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCCCTTAG 480
QY 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAGTACTTTCTGGGCGAGAGCCCGGCACT 540
Db 481 ACATCAACATCCAGGAACCTCGCTGGGACCAAGTACTTTCTGGGCGAGAGCCCGGCACT 540
QY 541 TTTTCACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGCACAAGCTTCTTC 600
Db 541 TTTTCACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGCACAAGCTTCTTC 600
QY 601 GGAACATCGTGAGAACTACAGGGCGGCGGTGGTGAGCCCCAGGGATCACCGAGGACCAGC 660
Db 601 GGAACATCGTGAGAA----- 616
QY 661 TGTGAGGGGCCAAGTATGTATGACTCCGCCCTTCATCCGAGACAGGGAGAGGTGG 720
Db 617 ----- 616
QY 721 TCCTGATTGGCCGATGTGAGGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGC 780
Db 617 ----- 616
QY 781 TCACATTTACAGGAAGACCCCAACCGTGTGTCTGCGAGTGGGTGAATCAGTCCCTCA 840
Db 617 -----CTACAGGAAGACCCCAACCGTGTGTCTGCGAGTGGGTGAATCAGTCCCTCA 669
QY 841 ATGCCATTTTAATACTACTCCAACCCGAGTGGTGACACATCCCATCACTGTGAGGCGAGCTGG 900
Db 670 ATGCCATTTTAATACTACTCCAACCCGAGTGGTGACACATCCCATCACTGTGAGGCGAGCTGG 729
QY 901 GGACAGCCTATGTGAGTGCCACCACTGGAGCTGTGGCCACGGCCCTGGACTCAAAATCCC 960
Db 730 GGACAGCCTATGTGAGTGCCACCACTGGAGCTGTGGCCACGGCCCTGGACTCAAAATCCC 789
QY 961 TCACCAAGCACCTGCCCCCTTGGTGGCGAGATTGTGCCCTTTGACAGCAGTGGCAGCTG 1020
Db 790 TCACCAAGCACCTGCCCCCTTGGTGGCGAGATTGTGCCCTTTGACAGCAGTGGCAGCTG 849
QY 1021 CCAACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGAGCTGAGGTGGCATCCCGGTGG 1080
Db 850 CCAACTGCATCAACATCCCCCTGATGAGGCGAGAGAGAGAGCTGAGGTGGCATCCCGGTGG 909
QY 1081 CTGATGAGGCGAGGTCAGAGGCTTGGCTACTCGGTGACTGCAAGCAAGGGAATCTTCC 1140
Db 910 CTGATGAGGCGAGGTCAGAGGCTTGGCTACTCGGTGACTGCAAGCAAGGGAATCTTCC 969
QY 1141 AGGTGGTGAATTCAGAATCTGATGGCGATTCTCTGCCATGGCCATCCACCACTGATCA 1200
Db 970 AGGTGGTGAATTCAGAATCTGATGGCGATTCTCTGCCATGGCCATCCACCACTGATCA 1029
QY 1201 TGGACACTCTGGAGAAGAAAGACTTCTCTGAAG 1232
Db 1030 TGGACACTCTGGAGAAGAAAGACTTCTCTGAAG 1061

QY	736	TGTCAGCCCGCAGGTGCCCATGAACATGACCATTACTGGCTGCATGCTCACATTTCTACAGA	795
Db	402	TGTCAGCCCGCAGGTGCCCATGAACATGACCATTACTGGCTGCATGCTCACATTTCTACAGA	461
QY	796	AGACCCCAACCGTGGTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTTGTTAACT	855
Db	462	CTACGCCCGGTGTGCTGTCTGGCAGTGGATTAAACCACTCCCTTCAATGCCCGTGTCAATT	521
QY	856	ACTCCAACCCGAGTGGTGACACTCCCATCACTGTGAGGCACTGGGGACAGCCCTATGTGA	915
Db	522	ACACCAACAGAAGTGGAGACGCAACCCCTCACTGTCAATGAGTTGGGAACACAGCTTAAGCTTT	581
QY	916	GTCGCCACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAATCCCTCAACCAAGCACCTGC	975
Db	582	CTGCAACAACCTGGTGCCGTAGCAACAGCTTAGAGCTCAATGCATTTGACCAAGCATGTCT	641
QY	976	CCCCCTTGGTCGGCAGATTTGTGCCCTTTGCAGCAGTGGCAGCTGCCAAGTGCATCAACA	1039
Db	642	CACCACTGATAGACGTTTTGTTCCCTTGTGCTGCCGTAGCTGCTGCTAATTGCATTAATA	701
QY	1036	TCCCCCTGATGAGGCAGACAGACAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAGGTC	1099
Db	702	TTCCATTATGAGGCCAAGGGAACCTCAAGTTGGCATTTCCCGTCACGGATGAGAATGGGA	761
QY	1096	AGAGGCTTGGCTACTCGGTGACTGCAAGCCAAAGCAGGGAATCTTCCAGGTGGTGAATTCAA	1155
Db	762	ACCGCTTGGGGAGTGGCGGAACGCTGCGAACAAGCCATCAGCAAGTTGTGTTGCCA	821
QY	1156	GAATCTGCATGGCGAATCTCTGCCATGGCCATCCCAACCACTGATTCATGACACTCTGGAGA	1215
Db	822	GGATTCTCATGGCAGGCCCTGGCATGGCCATCCCTCCATTCATTATGAACACTTTGGAAA	881
QY	1216	AGAAAGACTTCTGAAG 1232	
Db	882	AGAAAGCCTTTTGAAG 898	

```

RESULT 6
US-10-198-846-9767
; Sequence 9767, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: fastseq for windows version 4.0
; SEQ ID NO 9767
; LENGTH: 2978
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 2974, 2975, 2976, 2977, 2978
; OTHER INFORMATION: n - A,T,C or G
US-10-198-846-9767

```

Query Match	32.38;	Score 398.2;	DB 14;	Length 2978;
Best Local Similarity	69.18;	Pred. No. 3.7e-110;		
Matches 544;	Conservative	0;	Mismatches 243;	Indels 0;
				Gaps 0;

QY	446	AGCGATGGAAGCAAAATGGGTGATTTGCTTTAGACATCAACATCCAGAACCTCGCTG	505
Db	109	AGCGTCCGGGACCATGTCTTGGAGAACTACCACCAACATTAAACATCAAGGAACCTCGATG	168

QY		506	GGACCAAGTACTTTCCCTGGGCAGAGCCCCGCGCACCCTTTTTCACGTGTACTGATCCTCGAAA	565
Db		169	GGATCAAAAGCACATTTCATGTGACGAGGCCAATCATTTCTTCTACTGTAACTGACCCAGGAA	228
QY		566	TCTGCTGCTGTCCGGGCGACGCTGGAAGCTTCTCGAACATCTGTGCAGAATACTACAGGC	625
Db		229	CATTCTGTTAACCAACGACAACCTCGAGAGTGCGAGAAAAATAGTACATGATTACAGGCA	288
QY		626	CGGCGTGTGACCCCAGGAGATCACCGAGACCAGCTGTGGAGGGCCAAGTATGTATGA	685
Db		289	AGGAATGTTCCTCCTGGTCTTACAGAAATGAATGTGGAGAGCAAAGTACATCTATGA	348
QY		686	CTCCGCCCTTCATCCGGACACAGGGGAGAGAGGTGTCCTGATTTGGCCGCATGTACGCCCA	745
Db		349	TTCAAGCTTTTCATCCCTGACACTGCTGAGAGATGATTTTGATAGGAAGATGTACGCCCA	408
QY		746	GGTGCCCATGAACATGACCAATCCTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAAC	805
Db		409	GGTTC CATGAACATGACCAATCAGAGTTGTATGATGACGTTTTACAGGACTACGCCGGC	468
QY		806	CGTGTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCGATTTGTAATACTACTCCAACCG	865
Db		469	TGTGCTGTCTGGCAGTGAATTAACAGTCCCTTCAATGCCGTGCTCAATTACACCAACAG	528
QY		866	CAGTGTGACACTCCCAATCACTGTGAGGACGCTGGGGACAGCCTATGTGAGTGCCACCAC	925
Db		529	AAGTGGAGACGACCCCTCACTGTCAATGAGTTGGGAACAGCTTAGCTTCTGCAACAAC	588
QY		926	TGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGGT	985
Db		589	TGGTGGCGTAGCACACAGCTCTAGGACTCAATGCAATGACCAAGCATGTCTCACCACTGAT	648
QY		986	CGGCAGATTGTGCCCTTTGACAGCAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGAT	1045
Db		649	AGGACGTTTGTTCCTTCTGCTGCCGTAGCTGCTGCTAATTGCAATTAATATTCATTAAT	708
QY		1046	GAGGCAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCGAGTCCAGAGGCTTGG	1105
Db		709	GAGGCAAAAGGAACCTCAAGTTGGCATTTCCGTCACGAGTGAAGATGGGAACCGCTTGGG	768
QY		1106	CTACTCGTGACTGCAGCCACAGGGAATCTTCCAGTGTGATTTCAAGAAATCTGCAT	1165
Db		769	GGAGTCGGCAACGCTGGGAACAACAGCCATCACGCAAGTTGCTGTCAGAGATTTCTCAT	828
QY		1166	GGCGATTCTCTGCATGGCCATCCCAACCACTGATCATGAGACACTCTGGAGAGAAAGACTT	1225
Db		829	GGCAGCCCCCTGGCATGGCCATCCCTCCATTCATTATGAACACTTTGGAAAAAGAAAGCCTT	888
QY		1226	CCTGAAG 1232	
Db		889	TTTGAAG 895	

```

RESULT 7
US-10-014-338-1
; Sequence 1, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens

```



```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1121)..(1122)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1125)..(1126)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1132)..(1133)
; OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-1
```

Query Match 32.2%; Score 396.8; DB 14; Length 1134;
Best Local Similarity 69.8%; Pred. No. 7.3e-110;
Matches 536; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

```
QY 465 GGTGAATGCCCTTTAGACATCAATCCAGAACCTGCTGGACCAAGTACTTCTG 524
DB 17 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAGCACTTTCATT 76
QY 525 GGCAGAGCCCGGCACTTTTTCACCTGTTACTGATCTCGAAATCTGCTGTCGGGCA 584
DB 77 GGACGAGCCCAATCATTTCTTCACTGTACTGACCCAGAACATCTGTTAACCAACGAA 136
QY 585 CAGCTGGAAGCTTCTCGGAACATCTGCAAGACTACAGGGCCGGCTGTGACCCAGGG 644
DB 137 CAACTCGAGAGTGCAGAAAAATAGTACATGATTTACAGCAAGAAATGTTCTCCTGCT 196
QY 645 ATCACCAGAGACCAAGCTGTGAGGGCCCAAGATGTGTATGACTCCGCCCTCCATCCGGAC 704
DB 197 CTACAGAAATGAATGTGTGAGAGCAAGAACATCTATGATTCAGCTTTTCATCCTGAC 256
QY 705 ACAGGGAGAGAGTGTCTCTGATTTGCGCCGATGTCAAGCCAGGTGCCCATGAACATGACC 764
DB 257 ACTGGTGAAGATGATTTTGTATGAGAAAGATGTCAAGCCAGGTTCCCATGAACATGACC 316
QY 765 ATCACTGGCTGATGCTCAACATTTCTACAGAGAACCCCAACCCGCTGCTTCTGCGAGTGG 824
DB 317 ATCAGAGGTGTATGATGACGTTTACAGGACTACGCCGGCTGTCTGTTCTGCGAGTGG 376
QY 825 GTGAATCAGTCTTCAATGCCATTTGTAATCTCAACCCAGAGTGTGACACTCCCATC 884
DB 377 ATTAACCAAGTCTTCAATGCCGCTGCTCAATTAACCAACAGAAAGTGAGACGCCCTC 436
QY 885 ACTGTGAGGCACTGGGGACAGCCCTATGTGATGTCACCACTGGAGCTGGCCACGGCC 944
DB 437 ACTGTCAATGAGTTGGGAACAGCTTACGTTTGTCAACAACACTGTGCGCTAGCAACAGCT 496
QY 945 CTGGACTCAATCCCTCACCAGACACCTGCCCTTGTGCGGAGATTGTGCCCTTT 1004
DB 497 CTAGACTCAATGATGATGACCAAGCATGTCTCACCACTGATAGAGAGCTTGTCCCTTT 556
QY 1005 GCAGCAGTGGAGCTGCGCAACTGCATCAACATCCCTCTGATGAGGACAGAGAGCTGCAG 1064
DB 557 GCTGCCGTAGCTGCTGCTAATGCTAATTAATTCATTAATGAGGCAAGGAACTCAAA 616
QY 1065 GTGGCATCCCGTGGCTGATGAGGCAAGTCAAGGCTTGGCTACTCGGTGACTGCAGCC 1124
DB 617 GTTGGCATTCCTGCTCAAGGATGAGAAATGGAACCGCTTGGGGAGTGGCGAAAGCTGCG 676
QY 1125 AAGCAGGGAATCTTCAAGGTGTGATTTCAAGAAATCTGCATGGCGATTCCTGCGATGGCC 1184
DB 677 AAACAAGCCATCAAGCAAGTGTGCTGCTCAGAGATTCATGAGCCCTGCGCATGGCC 736
QY 1185 ATCCCACTGATGATGAGACACTGTGAGAGAAAGACTTCTCTGAAG 1232
DB 737 ATCCCTCATGATGATGAGAACTTTGAAAGAAAGAAAGCCCTTTTGAAG 784
```

RESULT 8
US-09-990-415A-5

```
; Sequence 5, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1567
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1015)
US-09-990-415A-5
```

Query Match 27.2%; Score 335.4; DB 10; Length 1567;
Best Local Similarity 64.5%; Pred. No. 3.3e-91;
Matches 501; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

```
QY 456 AGCAAAATGGTGAATTCCTTTAGACATCAATCCAGAACCTGCTGGACCAAGT 515
DB 41 AGCAAGATGAGGCTGACCTGTCTGCTTCAATCATGATGCCCGCTTGGAGACCAAGCC 100
QY 516 ACTTCTGCGGAGAGCCCGGCACTTTTCACTGTTACTGATCTCGAAATCTGCTGCTG 575
DB 101 ACCTTCTGCGGAGAGTGAAGCACTCTCTTAACATCAACGACCCCGCACTGTCTTTGTA 160
QY 576 TCCGGGACACCTGGAAGCTTCTCGAACAATCTGTGCAAGACTACAGGCGCGCTGTG 635
DB 161 TCTGAGCGGAGCTGAGCTGGCCAAAGTGTGAGAGAGACAGATGGGGTTGTG 220
QY 636 ACCCAGGATCAACCGAGAGACAGCTGTGAGGGCCAAATATGTATGACTCCGCTTC 695
DB 221 CCCCAGGACCAAGTGAAGAGCTGCTGTATGCCAAGAAGCTGATGACTCGGCTTC 280
QY 696 CATCCGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTCAAGCCAGTGGCCATG 755
DB 281 CACCCGACACTGGGAGAGAGATGATGATCGGGCGCATGTCTTCAGCTTCCCTGGC 340
QY 756 AACATGACCATCTGGCTGCTGCTCAATTTCTACAGAGAACCCCAACCGTGTGCTC 815
DB 341 GGCAATGATCAATACGGGCTTCAATGCTCCAGTTCTACAGAGATGCCGGGCTGATCTTC 400
QY 816 TGGCAGTGGTGAATCAAGTCTTCAATGCCATTTGTAATCTCAACCCAGTGTGAC 875
DB 401 TGGCAGTGGTGAACCAAGTCTTCAATGCCTTAGTCAACTACCAACAGAAATGCGGCT 460
QY 876 ACTCCCATCTGTGAGGCAAGTGGGAGCAAGCTATGTGAGTGCACCACTGGAGCTGTG 935
DB 461 TCCCCCATCATGATGAGGAGATGGCCCTTCTTCACTTCAAGCCACACCACTGCTGTG 520
QY 936 GCCACGGCCCTGGGACTCAAAATCCCTACCAAGCAAGTGGCCCTTGGTGGCAGATT 995
DB 521 GCCACGGCTGGGCAATGAACATGTGACAAAGAAAGCGCCGCTGTGGGCGCTGG 580
QY 996 GTGCCCTTGCAGCAGTGGCAGCTGCCAATCTCATCAACATCCCTGATGAGGACAGAG 1055
DB 581 GTGCCCTTGGCGCTGTGGCTGCGGCTAATGTGTCAATATCCCATGATGCGACAGAG 640
QY 1056 GAGCTGACAGTGGGCAATCCCGGTGGCTGATGAGGCAAGTCAAGAGGCTTGGCTACTCGGTG 1115
DB 641 GAGCTCATAAAGGAATCTGCGTGAAGGACAGGAATGAAATGAGATTGCTATCCCGG 700
QY 1116 ACTGACGCCAAGCAAGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCATTCCT 1175
DB 701 AGAGCTGGGCTAGGCAATCAACCAAGTATGATTTCTCGGATCAGCATGTCAAGCTCCT 760
QY 1176 GCCATGGCATCCCAACCACTGATCATGACACTGTGAGAGAAAGACTTCTCTGAAG 1232
```


Db 761 GGGATGATCTTGCTGCCAGTCATCATGGAAGGCTTGAGAAATTGCACCTTCATGCAG 817

RESULT 9
US-10-098-841-64

; Sequence 64, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098, 841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pl_fl_genes Version 1.0
; SEQ ID NO 64
; LENGTH: 2376
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (359)..(1327)
; US-10-098-841-64

Query Match 27.0%; Score 332.2; DB 13; Length 2376;
Best Local Similarity 64.2%; Pred. No. 3.5e-90;
Matches 499; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

QY 456 AGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGTGGACCAAGT 515
Db 353 AGCAAGATGAGGCTGACCTGTCTGCTTTAACATCATGATGCCCCCGTTGGACGCGC 412
QY 516 ACTTTCCTGGGAGAGCCGCGACCTTTTTCACCTGTTACTGATTCCTCGAATCTGCTGCTG 575
Db 413 ACCTTCCTGGGAGAGTGAAGCACTTCCTAACAATCAGGACCCCGGCACTGTCTTTGTA 472
QY 576 TCCGGGGACACAGCTGGAAGCTTCTCGGAACATCCTGCAGAACTACAGGGCCGCGTGTG 635
Db 473 TCTGAGCGGGAGCTGGACTGGGCCAAGGTGATGTTGAGAGAGCAGGATGGGGGTTGTG 532
QY 636 ACCCCAGGGATCACCGAGGACAGCTGTGGAGGGCCCAAGTATGTATGACTCCGCTTC 695
Db 533 CCCCCAGGCAACCAAGTGAAGCAGCTGTCTGTATGCCAAGAACTGTATGACTCGGCTTC 592
QY 696 CATCCGGACACAGGGGAGAGGTGCTCTGATTTGGCCGATGTACAGCCAGGTGCCATG 755
Db 593 CACCCGACACTGGGGAGAGATGAATGTATCGGGCGCATGTCTTTCAGCTTCCTGCGC 652
QY 756 AACATGACCATCACTGGCTGATGCTCACAATTTACAGAGAGACCCCAACCGTGTGTTC 815
Db 653 GGCATGATCATCAGCGGCTTCATGCTCCAGTTCTACAGAGAGATGCCGCGGTGATCTTC 712

QY 816 TGGCAGTGGGTGAATCAGTCCCTTCATGCCATTTGTTAACTACTCCAAACCGCAGTGTGAC 875
Db 713 TGGCAGTGGGTGAACACAGTCCCTTCATGCCATTTGTTAACTACTCCAAACAGGAATGCGCT 772
QY 876 ACTCCATCACTGTGAGGACAGTGGGACAGCCTATGTGAGTGCACCACTGAGCTGTG 935
Db 773 TCCCCACATCACTGACGAGCAGATGGCCCTTTCCTACTTACAGACCAACCACTGCTGTG 832
QY 936 GCCAGGGCCCTGGGACTCAAAATCCCTCACCAGACACCTGCCCCCTTGTCGGCAGATT 995
Db 833 GCCAGGCTGTGGGATGAACATGTGACAAAGAAAGCCGCCCTTGTTGGGCCCTGG 892
QY 996 GTGCCCTTTGCAGCAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGATGAGCAGAGA 1055
Db 893 GTGCCCTTTGCCGCTGTGGCTGGGCTTAACCTGTGTCAATATCCCATGATGCGACAGAG 952
QY 1056 GAGCTGCAGCTGGGATCCCGTGGCTGATGAGGACAGTCAAGGCTTGCTACTCGGTG 1115
Db 953 GAGCTCATAAAGGAATCTGCTGAGAGACAGGAATGAAATGAGATTGTCATTCCCGG 1012
QY 1116 ACTGCAGCCAAAGCAGGAATCTCCAGGTGTTGATTTCAAGAAATCTGCATGGCATTCCT 1175
Db 1013 AGAGCTGGGCGCATAGGCATCACCACAGTAGTATTCTCGATCACCATGTCAGCTCCT 1072
QY 1176 GCCATGCCCATCCCACTGATCATGAGACACTCTGAGAGAAAGACTTCCTGAAG 1232
Db 1073 GGGATGATCTTGCTGCCAGTCATCATGGAAGGCTTGAGAAATTCACCTTCATGCAG 1129

RESULT 10

US-09-822-849A-490
; Sequence 490, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822, 849A
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195, 582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 490
; LENGTH: 2458
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-822-849A-490

Query Match 26.8%; Score 330.6; DB 9; Length 2458;
Best Local Similarity 64.1%; Pred. No. 1.1e-89;
Matches 498; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

QY 456 AGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGAACCTCGCTGGACCAAGT 515
Db 75 AGCAAGATGAGGCTGACCTGTCTGCTTTAACAATCGATGCCCCCGTTGGACCAAGCC 134
QY 516 ACTTTCCTGGGAGAGCCCGGCACTTTTTCACCTGTTACTGATCTCGAAATCTGCTGCTG 575
Db 135 ACCTTCCTGGGAGAGTGAAGCACTTCTTAAACATCACGAGACCCCGCACTGTCTTGTGA 194
QY 576 TCCGGGGACACAGCTGGAAGCTTTCGGAACATCGTGAGAACTACAGGGCCGCGTGTG 635
Db 195 TCTGAGCGGAGAGCTGACTGGGCCAAGGTGATGTTGAGAGAGCAGATGGGGTGTG 254

```
QY 636 ACCCCAGGATCACCAGACCACTGTGAGGGCCCAATATGTATGACTCCCGCTTC 695
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 255 CCCCCAGGCACCCCAAGTGGAGCAGCTGCTGTATGCCAAAAGCTGTATGACTCGGCTTC 314
QY 696 CATCCGGACACAGGGGAGAGGTGCTCTGATTTGGCCGATGTACGCCAGGTGCCATG 755
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 315 CACCCCGACACTGGGAGAGATGAATGTATCATCGGCGCATGCTTTCCAGCTTCCTGCG 374
QY 756 AACATGACCATCACTGGCTGATGCTCACAATTTACAGAGAACCACAACCGTGTGTC 815
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 GGCATGATCATCAGCGGCTTCATGCTCCAGTTCTACAGACGATGCCGCGTGAATCTTC 434
QY 816 TGGCAGTGGTGAATCAGTCTTCAATGCCATGTTAATCTACTCCAACCGCAGTGTGAC 875
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 435 TGGCAGTGGTGAACCAAGTCTTCAATGCCCTTAGTCACTACACCAAGAAATGCGGCT 494
QY 876 ACTCCCATCACTGTGAGGCACTGGGAGACACCTATGTAGTGCACCACTGAGCTGTG 935
    . | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 495 TCCCCACATCATCAGGAGATGGCCCTTCTCTACTCTACACACCAACCACTGCTGTG 554
QY 936 GCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCTTGTGCGCAGATTT 995
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 555 GCCACGGCTGTGGGATGACATGTGACAAAGAAAGCCGCCCTTGTGGGCGGCTGTG 614
QY 996 GTGCCCTTTGCAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGCAGAGA 1055
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 615 GTGCCCTTTGCCGCTGTGGCTGGGCTTAATCTGTCAATATCCCATGATGCGACAGCAG 674
QY 1056 GAGCTGCAGGTGGGATCCCGTGGCTGATGAGGAGGCTGAGGCTTGGCTACTCGGTG 1115
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 675 GAGCTCATAAAGGAATCTGCGTGAAGACAGGAATGAAATGAGATTGTCATTCGCG 734
QY 1116 ACTGACGCCAAGCAGGAATCTTCCAGGTGTTGATTTCAAGAAATCTGCATGGCATTCCT 1175
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 735 AGAGCTGGCGGCATAGGCATCAACCAAGTAGTATTCTCGGATCACCATGTCAAGCTCT 794
QY 1176 GCCATGGCCATCCACCACTGATCATGAGACACTCTGAGAGAAGAACTTCTCTGAAG 1232
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 795 GGGATGATCTTGTCTGCCAGTCAATGGAAGGCTTGAGAAATTGCACTTCATGCAG 851
```

```
RESULT 11
US-10-014-338-3
; Sequence 3, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE AN
; FILE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (949)..(950)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (979)..(980)
; OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-3
```

Query Match 25.6%; Score 315; DB 14; Length 984;
Best Local Similarity 70.6%; Pred. No. 4.3e-85;

```
Matches 420; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY 465 GGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGCTGGGACCAAGTACTTCTG 524
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 GGAGAACTAACCAACATTAATCAATCAAGAACCTCGATGGGATCAAGCACTTCAAT 76
QY 525 GGCAGAGCCCGGCACTTTTTCACCTGTACTGATCTCGAAATCTGCTGCTCCGGGCA 584
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 GGACGAGCCAAATCATTTCTTCACTGTAACTGACCCCAAGAACATTTCTGTTAACCAAGAA 136
QY 585 CAGCTGGAAGCTTCTCGGAACATCGTGACAACTACAGGGCGGCGGTGTGACCCAGGG 644
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 CAACTCGAGAGTGCAGAGAAAATAGTACATGATTACAGGCAAGGAATTGTTCTCTGCT 196
QY 645 ATCAGCGAGACAGCTGTGAGAGGCCCAAGTATGTATGACTCCGCTTCATCCGAG 704
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 CTACAGAAAATGAATTTGTGAGAGCAAAAGTACATCTATGATTCAGCTTTTCATCTGAC 256
QY 705 ACAGGGGAGAAAGTGTCTGATTTGGCCGATGTGAGCCCAAGTGGTCCATGAACATGACC 764
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 ACTGTGAGAAAGATGATTTTGAATGAGAAAGATGTACGCCCAAGTCCCATGAACATGACC 316
QY 765 ATCACTGCTGCATGCTCACAATCTTACAGAGAAGAACCCCAACCGTGTGTTCTGCGAGTGG 824
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 ATCAGAGTGTATGATGACGTTTACAGAGACTACGCCGCGTGTGCTGTGCGCAGTGG 376
QY 825 GTGAATCAGTCTTCAATGCCATTTGTTAACTACTCCAAACCGCAGTGGTGACACTCCATC 884
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 ATTAACCAATCCTTCAATGCGCGTGTCAATTAACCAACAGAAAGTGAGACGACACCCCTC 436
QY 885 ACTGTAGAGCAGCTGGGAGACGCTATGTGAGTGCCACCACTGAGCTGTGCCACGCGC 944
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGTAACAACTGTGCGGTAGCAACAGCT 496
QY 945 CTGGACTCAATCCCTCACCAGCACTGCCCGCTTGTGGCAGATTTGTGCCCTTT 1004
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 497 CTAGACTCAATGCAATTGACCAAGCATGCTCACCACCTGATAGCAGGTTTGTTCCTTT 556
QY 1005 GCAGAGTGGCAGCTGCCAACTGCATCAACATCCCGCTGATGAGCAGAGAGAGC 1059
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 557 GCTGCCGTAGCTGCTGCTAATTGCAATTAATTCATTAAATGAGGCAAAAGCCATC 611
```

```
RESULT 12
US-10-206-901B-41
; Sequence 41, Application US/10206901B
; Publication No. US20030100540A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, ZHONGHUA
; APPLICANT: DUBOIS, RAYMOND
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES
; FILE REFERENCE: VBLT:012US
; CURRENT APPLICATION NUMBER: US/10/206,901B
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/308,370
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-901B-41
```

Query Match 22.1%; Score 272.8; DB 14; Length 670;
Best Local Similarity 71.6%; Pred. No. 2.4e-72;
Matches 358; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

```
QY 733 GCATGTACGCCAGGTGCCCATGACATGACCATCACTGGCTGCATGCTCACATTTCTACA 792
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 GAATGTACAGCCCAAGTTCCCATGAACATGACCATCACAGGTTCTATGATGACGTTTACA 63
QY 793 GGAAGACCCCAACCGTGTGTCTGCGCAGTGGGTGAATCAGTCCCTTCATGCCATTTGTTA 852
```

```

Db      64  GGACTACGCGCGCTGTGCTGTCTGCGAGTGAATTAACCAAGTCCCTCAATGCCCTGCTCA
QY      853  ACTACTCCAAACCGAGTGTGACACTCCCATCTGTGAGGAGCTGGGAGACGCTATG
Db      124  ATTACACCAAGAGTGAAGACGACCCCTCACTGATGAGTGGAGACGCTTACG
QY      913  TGAGTCCACCACTGAGAGTGTGGCCAGCGCCCTGGGACTCAATCCCTCAACCAAGCAC
Db      184  TTCTGCAACAACTGTGCGGTAGCAACAGCTTAGAGCTCAATGACATGACCAAGCATG
QY      973  TGCCCCCTGTGTGCGCAGATTTGTGCCCTTTGACAGCAGTGGCAGCTGCCAAGTCA
Db      244  TCTCACCACCTGATAGGACGTTTGTCCCTTTGCTGCCGTAGCTGCTGCTTAATTGCA
QY      1033  ACATCCCCCTGATGAGGACAGAGAGCTGCGAGTGGGCTATCCCGTGGCTGATGAGG
Db      304  ATATTCCATTAAATGAGGCAAGGGAAGTCAAGTTGGCATTTCCCTCAGCGATGAGA
QY      1093  GTGAGAGGCTGTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      364  GGAACCGCTGTGGGGAGTGGCGAAGCGTGGCAACCAAGCCATCAGCAGATGCTGCT
QY      1153  CAAGAATCTGATGCGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      424  CCAGATTCCTGATGCGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
QY      1213  AGAAGAAAGACTTCTGTAAG 1232
Db      484  AAAAGAAAGCCTTTTGAAG 503

```

RESULT 13

```

US-09-918-995-23576
; Sequence 23576, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23576
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23576

```

Query Match 17.2%; Score 212.4; DB 11; Length 489;

Best Local Similarity 69.6%; Pred. No. 4.3e-54;

Matches 288; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

```

QY      446  AGCGATGAAAGCAAAATGGGTAATGCTTTAGACATCAACATCCAGGAACCTCGCTG
Db      76   AGCGTCCGGGACCAATGCTGTAGAACTACCAACAACATTAACATCAAGAACCTCGATG
QY      506  GGACCAAAAGTACTTCTCTGGGAGAGCCCGGCACTTTTCACTGTTACTGATCTCGAAA
Db      136  GGATCAAAAGCACTTCTCTGGAGAGAGCAATCTTCTCACTGTAAGTACCCAGGAA
QY      566  TCTGCTGCTGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGAGAACTACAGGCG
Db      196  CATCTGTTAACCAAGCAAACTCGAGAGTGGAGAAAATAGTACATGATTACAGGCA
QY      626  CGGCGTGTGACCCAGGATCAGGAGACCAAGCTGTGAGGGCAAGTATGTATGA
Db      256  AGGAATGTCTCTCTGCTCTTACAGAAAATGAATGTGTGAGAGCAAAAGTACATCTATGA

```

```

QY      686  CTCGCGCTTCATCCGGACACAGGGGAGAGAGGTGCTCTGATTTGCCCATGTACGCCA
Db      316  TTCAGCTTTTCATCCCTGACACTGTGTGAGAGATGATTTGATAGAGAAATGTACGCCA
QY      746  GGTGCCATGAACATGACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Db      376  GGTTCATGAACATGACATCAAGGTTGATGATGAGCTTTTACAGAGACTACGCCGCG
QY      806  CGTGTGTTCTGCGAGTGGGTGAATCAGTCTTCAATGCCATTTGTTACTACTC
Db      436  TGTGCTGTCTGCGAGTGAATTAACCAAGTCTTCAATGCCGCTGCTCAATTACAC

```

RESULT 14

```

US-09-783-590-2110
; Sequence 2110, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2110
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (35)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (82)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (159)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (172)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (177)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (181)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (212)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (225)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (233)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (262)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (275)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (281)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 13:08:36 ; Search time 2505 Seconds
(without alignments)
11953.339 Million cell updates/sec

Title: US-09-990-415A-1
Perfect score: 1232
Sequence: 1 cccttagcgccagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835.4	67.8	1201	13 BX376574	BX376574 BX376574
2	824.6	66.9	914	9 AL521267	AL521267 AL521267
3	816.4	66.3	869	13 BQ223306	BQ223306 AGENCOURT
4	798.8	64.8	998	12 BM473966	BM473966 AGENCOURT

5	798.6	64.8	995	13 BX355800	BX355800 BX355800
6	795.8	64.6	975	13 BQ933752	BQ933752 AGENCOURT
7	795.4	64.6	1201	13 BX334089	BX334089 BX334089
8	773.8	62.8	777	12 BI908654	BI908654 603066774
9	767.2	62.3	840	9 AL522228	AL522228 AL522228
10	759.6	61.7	811	12 BI818854	BI818854 603037323
11	759.4	61.6	1076	12 BM925231	BM925231 AGENCOURT
12	738.4	59.9	1014	13 BQ717821	BQ717821 AGENCOURT
13	727.2	59.0	862	13 BQ425697	BQ425697 AGENCOURT
14	713	57.9	963	13 BQ959114	BQ959114 AGENCOURT
15	711.6	57.8	1201	13 BX393485	BX393485 BX393485
16	705.8	57.3	1040	12 BM548674	BM548674 AGENCOURT
17	700.4	56.9	961	12 BG775086	BG775086 602650054
18	700	56.8	755	12 BI517760	BI517760 603042003
19	691.2	56.1	888	12 BM018524	BM018524 603646357
20	690.8	56.1	758	10 BF346247	BF346247 602018434
21	685.2	55.6	1008	10 BF569711	BF569711 602186237
22	680	55.2	1071	10 BF569784	BF569784 602185719
23	678.4	55.1	868	10 BE780804	BE780804 601469693
24	670.4	54.4	870	12 BM047107	BM047107 603627441
25	658.2	53.4	821	12 BI819328	BI819328 603037779
26	648.8	52.7	898	14 CB203758	CB203758 AGENCOURT
27	646.8	52.5	2682	11 AK034514	AK034514 Mus muscu
28	617.8	50.1	634	9 AL701236	AL701236 DKF2p6860
29	614	49.8	934	13 BU506082	BU506082 AGENCOURT
30	583.2	47.3	764	14 CA317392	CA317392 UI-M-FW0-
31	581.6	47.2	731	14 CA317398	CA317398 UI-M-FW0-
32	577.6	46.9	746	12 BI752914	BI752914 603028557
33	576	46.8	715	14 CA316261	CA316261 UI-M-FW0-
34	562.8	45.7	1094	10 BG331175	BG331175 602431879
35	552.2	44.8	558	9 AL602426	AL602426 DKF2p686J
36	548	44.5	727	14 CD354095	CD354095 UI-M-GM0-
37	534.2	43.4	666	14 CA319631	CA319631 UI-M-FW0-
38	527.4	42.8	529	10 BE256336	BE256336 601111229
39	508.4	41.3	676	12 BM944910	BM944910 UI-M-EH0P
40	505	41.0	655	10 BB626574	BB626574 BB626574
41	497.6	40.4	782	14 CB234292	CB234292 AGENCOURT
42	487.2	39.5	637	14 CA319259	CA319259 UI-M-FW0-
43	485.4	39.4	515	9 AW601313	AW601313 PMO-BT034
44	483.2	39.2	979	10 BG245508	BG245508 602355067
45	477.4	38.7	614	10 BB659747	BB659747 BB659747

ALIGNMENTS

RESULT 1
BX376574
LOCUS 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX376574 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
CDNA clone CS0DD001YF06 5-PRIME, mRNA sequence.
ACCESSION BX376574
VERSION BX376574.1 GI:30448513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5549.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD001DC03QPI&cluster=5549.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD001DC03QP1.										
FEATURES	location/Qualifiers									
source	1..1201									
	/organism="Homo sapiens"									
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="CS0DD001YF06"									
	/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"									
	/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"									
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized."									
BASE COUNT	268 a	306 c	347 g	250 t	30 others					
ORIGIN										
Query Match	67.8%;	Score 835.4;	DB 13;	length 1201;						
Best Local Similarity	98.5%;	Pred. No. 4.6e-196;								
Matches	842;	Conservative	1;	Mismatches	12;	Indels	0;	Gaps	0;	
QY	378	GTGATGGCTGGAGGGCCGGCGCGGACAGCGGAGGACAGAGAGAAAGCGGTCTGAGAG	437							
Db	41	GTACCGGTCGGAAATTCGGGGATCGACAGCGAGGCGAGAGAGAAAGCGGTCTGAGAG	100							
QY	438	CTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAA	497							
Db	101	CTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAA	160							
QY	498	CCTCGCTGGAGACCAAGTACTTCTCTGGGACAGAGCCCGGCACTTTTCACTGTACTGAT	557							
Db	161	CCTCGCTGGAGACCAAGTACTTCTCTGGGACAGAGCCCGGCACTTTTCACTGTACTGAT	220							
QY	558	CCTCGAATCTGCTGCTTCCGGGGGACACAGCTGGAAGCTTCTCGGAACATCGTCAGAAC	617							
Db	221	CCTCGAATCTGCTGCTTCCGGGGGACACAGCTGGAAGCTTCTCGGAACATCGTCAGAAC	280							
QY	618	TACAGGGCCGGCGTGGTGACCCACAGGATCACCGAGCACAGCTGTGGAGGGCCAAGTAT	677							
Db	281	TACAGGGCCGGCGTGGTGACCCACAGGATCACCGAGCACAGCTGTGGAGGGCCAAGTAT	340							
QY	678	GTGTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAGGTGCTGATTGGCCGATG	737							
Db	341	GTGTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAGGTGCTGATTGGCCGATG	400							
QY	738	TCAGCCGAGTGCCCATGAACATGACCATCTGCTGCATGCTCACATTCTACAGGAAG	797							
Db	401	TYAGCCGAGTGCCCATGAACATGACCATCTGCTGCATGCTCACATTCTACAGGAAG	460							
QY	798	ACCCCAACCGTGTCTTGCGCAGTGGGTGAATCAGTCTTCAATGCCATTTGTAACCTAC	857							
Db	461	ACCCCAACCGTGTGTCTTGCGCAGTGGGTGAATCAGTCTTCAATGCCATTTGTAACCTAC	520							
QY	858	TCCAACCGCAGTGTGTGACACTCCCATCACTGTGAGGCAGCTGGGGACAGCCCTATGTAGT	917							
Db	521	TCCAACCGCAGTGTGTGACACTCCCATCACTGTGAGGCAGCTGGGGACAGCCCTATGTAGT	580							
QY	918	GCCACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCC	977							
Db	581	GCCACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCC	640							
QY	978	CCCTTGGTGGCAGATTTGTGCCCTTTGACAGCAGTGGCAGCTGCGCAACTGCATCAACATC	1037							
Db	641	CCCTTGGTGGCAGATTTGTGCCCTTTGACAGCAGTGGCAGCTGCGCAACTGCATCAACATC	700							
QY	1038	CCCTTGATGAGGCGAGAGAGAGCTGCAGGTGGGCAATCCCGGTGGCTGATGAGGCGAGTCAAG	1097							
Db	701	CCCTTGATGAGGCGAGAGAGAGCTGCAGGTGGGCAATCCCGGTGGCTGATGAGGCGAGTCAAG	760							
QY	1098	AGGCTTGGCTACTTCGGTGAATGACAGCCAAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGA	1157							
Db	761	AGGCTTGGCTACTTCGGTGAATGACAGCCAAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGA	820							
QY	1158	ATCTGCATGGCGATTCTGCCATGGCCATCCACCACTGATCATGGACACTCTGGAGAAG	1217							

Db	821	ATCTGCATGGCGATTCCTGCCATGGCCATCCACCACCTGATCATGACACTCTGGAGAAG	880
QY	1218	AAAGACTTCCTGAAG	1232
Db	881	AAAGACTTCCTGAAG	895
RESULT 2			
AL521267		914 bp	mRNA
LOCUS			linear
DEFINITION	AL521267 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens		EST 22-MAY-2003
ACCESSION	AL521267		
VERSION	AL521267.2		GI:31039573
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 914)		
JOURNAL	L1,W.B., Gruber,C., Jessee,J. and Polayes,D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished		
	On Feb 13, 2001 this sequence version replaced gi:12784760.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by life technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 5549.f For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		
	cgl-bin/cluster.cgi?seq=CS0DB001CB06QPl&cluster=5549.f. Contact :		
	Feng Liang Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
	Paraday Avenue Genoscope sequence ID : CS0DB001CB06QPl.		
FEATURES			
Source	location/Qualifiers		
	1. 914		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DB001YD11"		
	/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"		
	/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-oligo(dT)		
	primer. Five prime end enriched, double-strand cDNA was		
	digested with Not I and cloned into the Not I and EcoR V		
	sites of the pCMVSPORT 6 vector. Library was normalized."		
BASE COUNT	179 a	265 c	277 g
ORIGIN		193 t	
Query Match	66.98;	Score 824.6;	DB 9;
Best Local Similarity	99.3%;	Pred. No. 1.9e-193;	
Matches	849;	Conservative	0; Mismatches 4; Indels 2; Gaps 2;
QY	43	CAGCGAGTCTTTACCCACCCAGTTCGTGTTCTGACGCGCTAGCTATTCGCAAAATT	102
Db	62	CCGGATGTTCTTACCCACCCAGTTCGTGTTCTGACGCGCTAGCTATTCGCAAAATT	121
QY	103	TAGGGCTTGGGTCGTGCTGTTCCTCCGCGCTCGAACCACTTCTCTGAGCCGAGCC	162
Db	122	TAGGGCTTGGGTCGTGCTGTTCCTCCGCGCTCGAACCACTTCTCTGAGCCGAGCC	181
QY	163	AGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGGGCACCCCTTGAGGCTCCGTGAG	222
Db	182	AGCTACCGGGCTCCTGGAATTGCCACCCCTCCCTGGGCACCCCTTGAGGCTCCGTGAG	241
QY	223	GGACGTACAGGGGAGAGCGGGAGCGTGAAGCTGAGTTTCTGACAGCGCTGCTGTGTGG	282
Db	242	GGACGTACAGGGGAGAGCGGGAGCGTGAAGCTGAGTTTCTGACAGCGCTGCTGTGTGG	301
QY	283	TGGCTGGGTTCTGCCAATCCCCGTGCCCCACCGGGGTGGCGCGCGGAGAGCTCCTGCCC	342

Db 302 TGGCTGGGTTCTGCCAATCCCGTCCACCGGGTGGCGCGCGGAGAGCTCTGCC 361

QY 343 CTCCTGCTGTCGGCGCTACGCGTGCAGTCCCGCTGATGCTGGAGGCGCGCGG 402

Db 362 CTCCTGCTGTCGGCGCTACGCGTGCAGTCCCGCTGATGCTGGAGGCGCGCGG 421

QY 403 GACAGCGGAGGAGAGAGAGCGCGTCTGAGAGCTTCAGAGCGATGGAAGCAAAA 462

Db 422 GACAGCGGAGGAGAGAGAGCGCGTCTGAGAGCTTCAGAGCGATGGAAGCAAAA 481

QY 463 TGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCAAGTACTTCC 522

Db 482 TGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCAAGTACTTCC 541

QY 523 TGGGACAGAGCCCGGACCTTTTCACTGTTACTGATCCTCGAAATCTGCTGTCGGGG 582

Db 542 TGGGACAGAG-CCGGACCTTTTCACTGTTACTGATCCTCGAAATCTGCTGTCGGGG 600

QY 583 CACAGCTGGAAGCTTCTCGAATCATGTCAGAACTACAGGCGCGCTGGTGAACCCAG 642

Db 601 CACAGCTGGAAGCTTCTCGAATCATGTCAGAACTACAGGCGCGCTGGTGAACCCAG 660

QY 643 GGATCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 702

Db 661 GGATCACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

QY 703 ACACAGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 762

Db 721 ACACAGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

QY 763 CCATCACTGGCTGCATGCTCACAATTTCTACAGGAAGAGGAGGAGGAGGAGGAG 822

Db 781 CCATCACTGGCTGCATGCTCACAATTTCTACAGGAAGAGGAGGAGGAGGAGGAG 839

QY 823 GGGTGAATCAGTCTCTCAATGCCATGTTAACTACTCCAAACGAGTGGTGAACCTCCA 882

Db 840 GGGTGAATCAGTCTCTCAATGCCATGTTAACTACTCCAAACGAGTGGTGAACCTCCA 899

QY 883 TCACTGTGAGGAGC 897

Db 900 TCACTGTGAGGAGC 914

RESULT 3
BQ223306 869 bp mRNA linear EST 02-MAY-2002
LOCUS BQ223306
DEFINITION AGENCOURT_7518440 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6066572
5', mRNA sequence.
ACCESSION BQ223306
VERSION BQ223306.1 GI:20404706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13344 row: k column: 21
High quality sequence stop: 688.

FEATURES
source 1..869

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6066572"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 193 a 244 c 255 g 175 t 2 others
ORIGIN

Query Match 66.3%; Score 816.4; DB 13; Length 869;
Best Local Similarity 99.8%; Pred. No. 2e-191;
Matches 817; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 414 CAGAGGAGAGGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTG 473

Db 1 CAGAGGAGAGGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTG 60

QY 474 CCTTAGACATCAACATCCAGAGACCTCGCTGGAGCCAAAGTACTTCTCGGACAGACC 533

Db 61 CCTTAGACATCAACATCCAGAGACCTCGCTGGAGCCAAAGTACTTCTCGGACAGACC 120

QY 534 CCGCACTTTTCACTGTTACTGATCTCGAATCTGCTGCTGCTGGGACACAGCTGGAA 593

Db 121 CCGCACTTTTCACTGTTACTGATCTCGAATCTGCTGCTGCTGGGACACAGCTGGAA 180

QY 594 GCTTTCGGAACATCGTGCAGAACTACAGGCGCGGCGTGGTGAACCCAGGATCACCGAG 653

Db 181 GCTTTCGGAACATCGTGCAGAACTACAGGCGCGGCGTGGTGAACCCAGGATCACCGAG 240

QY 654 GACCAGCTGTGAGGCGCAAGTATGTATGACTCCGCCCTTCATCCGGACACAGGGAG 713

Db 241 GACCAGCTGTGAGGCGCAAGTATGTATGACTCCGCCCTTCATCCGGACACAGGGAG 300

QY 714 AAGTGTCTCTGATTGGCCGATGTCAGCCAGGTGGCCATGACATGACATCACTGGC 773

Db 301 AAGTGTCTCTGATTGGCCGATGTCAGCCAGGTGGCCATGACATGACATCACTGGC 360

QY 774 TGCATGCTCACAATTTCTACAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 833

Db 361 TGCATGCTCACAATTTCTACAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

QY 834 TCCCTCAATGCCATGTTAACTACTCCAAACGAGTGGTGAACCTCCATCACTGTGAGG 893

Db 421 TCCCTCAATGCCATGTTAACTACTCCAAACGAGTGGTGAACCTCCATCACTGTGAGG 480

QY 894 CAGCTGGGGACAGCCTATGTGAGTGCACCACTGGAGCTGTGGCCAGGGCCCTGGAGCTC 953

Db 481 CAGCTGGGGACAGCCTATGTGAGTGCACCACTGGAGCTGTGGCCAGGGCCCTGGAGCTC 540

QY 954 AAATCCTTCACCAAGCACCTGCCCCCTTGTGCGGAGATTTGCCCTTTGCAGCAGTG 1013

Db 541 AAATCCTTCACCAAGCACCTGCCCCCTTGTGCGGAGATTTGCCCTTTGCAGCAGTG 600

QY 1014 GCAGCTGCCAACTGCATCAACATCCCTGTATGAGGACAGAGAGAGCTGCAGGTGGGATC 1073

Db 601 GCAGCTGCCAACTGCATCAACATCCCTGTATGAGGACAGAGAGAGCTGCAGGTGGGATC 660

QY 1074 CCGTGGCTGTATGAGGACAGGCTTGCTACTCGGTGACTGCAGCCAGAGAGGA 1133

Db 661 CCGTGGCTGTATGAGGACAGGCTTGCTACTCGGTGACTGCAGCCAGAGAGGA 720

QY 1134 ATCTTCAGGTGTGATTTCAAGAACTGCATGGCGATTCTCGCCATGGCCATCCACCA 1193

Db 721 ATCTTCAGGTGTGATTTCAAGAACTGCATGGCGATTCTCGCCATGGCCATCCACCA 780

QY 1194 CTGATCATGACACTCTGGAGAGAAAGACTTCTCTGAAG 1232

Db	781	CTGATCATGGACACTCTTGAGAGAAAGACTTCTGAAG	819
RESULT 4			
LOCUS	BM473966	998 bp	mRNA
DEFINITION	AGENCOURT_6489827 NIH_MGC_71	Homo sapiens	CDNA clone IMAGE:5521420
ACCESSION	BM473966		
VERSION	5'	mrna sequence.	
KEYWORDS	BM473966.1	GI:18523008	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 998)		
TITLE	NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM12187	row: a	column: 05
	High quality sequence stop: 652.		
FEATURES	location/Qualifiers		
source	1. 998		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5521420"		
	/tissue_type="leiomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_71"		
	/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."		
BASE COUNT	228 a	292 c	279 g 198 t 1 others
ORIGIN			
Query Match	64.8%;	Score 798.8;	DB 12; Length 998;
Best Local Similarity	99.4%;	Pred. No. 4.9e-187;	
Matches 822;	Conservative 0;	Mismatches 3;	Indels 2; Gaps 2;
QY	408	CGGAGCGAGAGAGGAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGT	467
Db	1	CGGAGCGAGAGAGGAGCGGTTCTGAGAGCTTCAGAGAGCTATGGAAGCAAAATGGGT	60
QY	468	GAATGCTTTAGACATCAACATCCAGGAACCTCGCTGGACCAAGTACTTTCCTGGGC	527
Db	61	GAATGCTTTAGACATCAACATCCAGGAACCTCGCTGGACCAAGTACTTTCCTGGGC	120
QY	528	AGAGCCCGGACACTTTTTCACACTGTTACTGATCCCTGAAATCTGCTGCTGCCGGGCACAG	587
Db	121	AGAGCCCGGACACTTTTTCACACTGTTACTGATCCCTGAAATCTGCTGCTGCCGGGCACAG	180
QY	588	CTGAGAGCTTCTCGGAACATCGTGCAGAACTACAGAGGCGCGCTGCTGAGACCCAGGATC	647
Db	181	CTGAGAGCTTCTCGGAACATCGTGCAGAACTACAGAGGCGCGCTGCTGAGACCCAGGATC	240
QY	648	ACCGAGGACCAAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCGAGACACA	707
Db	241	ACCGAGGACCAAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCGAGACACA	300
QY	708	GGGAGGAAGGTGCTCTGATTTGGCCGCATGTACAGCCAGGTGCGCCATGAACATGACCATC	767
Db	301	GGGAGGAAGGTGCTCTGATTTGGCCGCATGTACAGCCAGGTGCGCCATGAACATGACCATC	360

QY	768	ACTGGCTGCATGCTTCACATTTCTACAGGAAGACCCCAACCGTGTGTTCTGCAGTGGGTG	827
Db	361	ACTGGCTGCATGCTTCACATTTCTACAGGAAGACCCCAACCGTGTGTTCTGCAGTGGGTG	420
QY	828	AATCAGTCCCTTCATGCCATGCTTGTTAACCTACCCCAACCGCAGTGTGACACTCCCATCACT	887
Db	421	AATCAGTCCCTTCATGCCATGCTTGTTAACCTACCCCAACCGCAGTGTGACACTCCCATCACT	480
QY	888	GTGAGGCAGCTGGGGACAGCCCTATGTGAGTGGCCACCACCTGAGCTGTGGCCACGGCCCTG	947
Db	481	GTGAGGCAGCTGGGGACAGCCCTATGTGAGTGGCCACCACCTGAGCTGTGGCCACGGCCCTG	540
QY	948	GGACTCAAAATCCCTTCACCAAGCACCTGCCCCCTTGGTCGGCAGATTTGGCCCTTTGCA	1007
Db	541	GGACTCAAAATCCCTTCACCAAGCACCTGCCCCCTTGGTCGGCAGATTTGGCCCTTTGCA	600
QY	1008	GCAGTGGCAGCTGCCCAACTGCATCAACAATCCCCCTGATGAGGCAGAGAGAGCTGCAGGTG	1067
Db	601	GCAGTGGCAGCTGCCCAACTGCATCAACAATCCCCCTGATGAGGCAGAGAGAGCTGCAGGTG	660
QY	1068	GGCATCCCGGTGGCTGATGAGGCAGGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCAAAG	1127
Db	661	GGCATCCCGGTGGCTGATGAGGCAGGTCAAGAGGCTTGGCTACTCGGTGACTGCAGCCAAAG	720
QY	1128	CA-GGGAATCTTCCAGGT-GGTGATTTCAAGAATCTGCATGGCGAATTCCTGCCATGGCCA	1185
Db	721	CAGGGGAATCTTCCAGGTGGGTGATTTCAAGAATCTGCATGGCGAATTCCTGCCATGNCNA	780
QY	1186	TCCCACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCCCTGAAG	1232
Db	781	TCCCACCACTGATCATGACACTCTGAGAGAAGAAAGACTTCCCTGAAG	827

RESULT 5	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BX355800	BX355800	995 bp mRNA linear EST 05-MAY-2003	BX355800	BX355800	EST.	Homo sapiens (human)	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 995)	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished
			BX355800.1	GI:30384010				Contact: Genoscope				
								Genoscope - Centre National de Sequencage				
								BP 191 91006 EVRY cedex - France				
								Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
								Library was constructed by Life Technologies, a division of				
								Invitrogen. This sequence belongs to sequence cluster 5549.f For				
								more information about this cluster, see				
								http://www.genoscope.cns.fr/				
								cgi-bin/cluster.cgi?seq=CSODI004BB03QPlcluster=5549.f. Contact :				
								Feng Liang Email : fliang@lifetech.com URL :				
								http://fulllength.invitrogen.com/Invitrogen Corporation 1600				
								Faraday Avenue Genoscope sequence ID : CSODI004BB03QPl.				
								Location/Qualifiers				
								1. 995				
								/organism="Homo sapiens"				
								/mol_type="mRNA"				
								/db_xref="taxon:9606"				
								/clone="CSODI004YC06"				
								/tissue_type="PLACENTA COT 25-NORMALIZED"				
								/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"				
								/note="1st strand cDNA was primed with a NotI-oligo(dT)				
								primer. Five prime end enriched, double-strand cDNA was				
								digested with Not I and cloned into the Not I and EcoR V				
								sites of the pCMVSPORT 6 vector. Library was normalized."				

BASE COUNT	226 a	281 c	281 g	193 t	14 others
ORIGIN					
Query Match	64.8%; Score 798.6; DB 13; Length 995;				
Best Local Similarity	98.5%; Pred. No. 5.5e-187;				
Matches 795; Conservative	9; Mismatches 3; Indels 0; Gaps 0;				
QY	426	CGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCTTTAGACATC	485		
Db	67	CGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCTTTAGACATC	126		
QY	486	AACATCCAGGAACCTCGCTGGGACCAAGTACTTTCTGGGGCAGAGCCCGGCACTTTTTC	545		
Db	127	AACATCCAGGAACCTCGCTGGGACCAAGTACTTTCTGGGGCAGAGCCCGGCACTTTTTC	186		
QY	546	ACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGACACAGCTGGAAGCTTCTCGGAAC	605		
Db	187	ACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGACACAGCTGGAAGCTTCTCGGAAC	246		
QY	606	ATCGTCAGAACTACAGGGCCGGCGTGTGACCCAGGGATCACCGAGACCAAGCTGTGG	665		
Db	247	ATCGTCAGAACTACAGGGCCGGCGTGTGACCCAGGGATCACCGAGACCAAGCTGTGG	306		
QY	666	AGGCCCAAGTATGTGTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAAGGTGTCTG	725		
Db	307	AGGCCCAAGTATGTGTATGACTCCGCCCTTCCATCCGGACACAGGGGAGAAGGTGTCTG	366		
QY	726	ATTGGCCGATGTCCAGCCAGTGGCCATGAACATGACCATCAGTGGCTGCATGCTCACA	785		
Db	367	ATTGGCCGATGTCCAGCCAGTGGCCATGAACATGACCATCAGTGGCTGCATGCTCACA	426		
QY	786	TTCTACAGGAAGCCCAACCGTGTGTCTGGCAGTGGGTGAATCAGTCTTCAATGCC	845		
Db	427	TTCTACAGGAAGCCCAACCGTGTGTCTGGCAGTGGGTGAATCAGTCTTCAATGCC	486		
QY	846	ATTGTTAACTACTCCCAACCGGAGTGTGACACTCCCATCAGTGTGAGGCAAGCTGGGACA	905		
Db	487	ATTGTTAACTACTCCCAACCGGAGTGTGACACTCCCATCAGTGTGAGGCAAGCTGGGACA	546		
QY	906	GCCTATGTGAGTGCACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCAC	965		
Db	547	GCCTATGTGAGTGCACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAATCCCTCAC	606		
QY	966	AAGCACCTGCCCCCTTGGTCGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCTGCCAAC	1025		
Db	607	AAGCACCTGCCCCCTTGGTCGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCTGCCAAC	666		
QY	1026	TGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGAGTGGGCAATCCCGTGCTGAT	1085		
Db	667	TGCATCAACATCCCCCTGATGAGGCGAGAGAGAGCTGAGTGGGCAATCCCGTGCTGAT	726		
QY	1086	GAGGCAAGTCAGAGGCTTGGCTACTCGGTGACTGACGCCAAGCAGGGAATCTTCCAGTG	1145		
Db	727	GAGGCAAGTCAGAGGCTTGGCTACTCGGTGACTGACGCCAAGCAGGGAATCTTCCAGTG	786		
QY	1146	GTGATTTCAAGATCTGCATGGCGATTCTCGCCATGGCCATCCCACTGATCATGAGAC	1205		
Db	787	GTGATTTCAAGATCTGCATGGCGATTCTCGCCATGGCCATCCCACTGATCATGAGAC	846		
QY	1206	ACTCTGAGAGAAGAAAGACTTCTCTGAAG	1232		
Db	847	ACTCTGAGAGAAGAAAGACTTCTCTGAAG	873		
RESULT 6					
BQ933752					
LOCUS	BQ933752	975 bp	mRNA	linear	EST 21-AUG-2002
DEFINITION	AGENCOURT_8837545 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6424845				
ACCESSION	BQ933752				
VERSION	BQ933752.1 GI:22349135				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				

ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
AUTHORS	1 (bases 1 to 975)									
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .									
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished									
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: DCTD/DTP/Gazdar CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LCM2605 row: k column: 22 High quality sequence stop: 646.									
FEATURES	source 1..975 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6424845" /tissue_type="large cell carcinoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_18" /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."									
BASE COUNT	225	a	277	c	267	g	199	t	7	others
ORIGIN										
Query Match	64.6%; Score 795.8; DB 13; Length 975;									
Best Local Similarity	99.1%; Pred. No. 2.7e-186;									
Matches 800; Conservative	0; Mismatches 7; Indels 0; Gaps 0;									
QY	421	GAAAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTCCTTTAG	480							
Db	1	GAAAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTCCTTTAG	60							
QY	481	ACATCAACATCCAGGAACCTCGCTGGGACCAAGTACTTCTCGGCGAGAGCCGGCACT	540							
Db	61	ACATCAACATCCAGGAACCTCGCTGGGACCAAGTACTTCTCGGCGAGAGCCGGCACT	120							
QY	541	TTTTCACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGCACAGCTGGAAGCTTCTC	600							
Db	121	TTTTCACTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGCACAGCTGGAAGCTTCTC	180							
QY	601	GGAACATCTGCAGAACTACAGGGCCGGCTGTGTGACCCCAAGGATCACCGAGACCAGC	660							
Db	181	GGAACATCTGCAGAACTACAGGGCCGGCTGTGTGACCCCAAGGATCACCGAGACCAGC	240							
QY	661	TGTGAGGGCCAAAGTATGTATGACTCCGCTTCCATCCGAGACAGGGGAGAAGTGG	720							
Db	241	TGTGAGGGCCAAAGTATGTATGACTCCGCTTCCATCCGAGACAGGGGAGAAGTGG	300							
QY	721	TCCTGATGGCCGCATGTACAGCCCAAGTCCCATGAACATGACATCACTGGCTGCATGC	780							
Db	301	TCCTGATGGCCGCATGTACAGCCCAAGTCCCATGAACATGACATCACTGGCTGCATGC	360							
QY	781	TCACATTTACAGGAAGACCCCAACCGTGTCTTGCGAGTGGGTGAATCAGTCTTCA	840							
Db	361	TCACATTTACAGGAAGACCCCAACCGTGTCTTGCGAGTGGGTGAATCAGTCTTCA	420							
QY	841	ATGCCATGTTAATCTACTCCAACCGGATGTGTGACACTCCCATCACTGTGAGGCAAGCTGG	900							
Db	421	ATGCCATGTTAATCTACTCCAACCGGATGTGTGACACTCCCATCACTGTGAGGCAAGCTGG	480							

```
QY      901 GGACAGCCTATGTAGTGCCACCACTGGAGCTGTGCCACGGCCCTGGGACTCAAAATCCC 960
      |||
Db      481 GGACAGCCTATGTAGTGCCACCACTGGAGCTGTGCCACGGCCCTGGGACTCAAAATCCC 540
QY      961 TCACCAAGCACCCTGCCCCCTTGGTGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCTG 1020
      |||
Db      541 TCACCAAGCACCCTGCCCCCTTGGTGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCTG 600
QY      1021 CCAACTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGGTGGCATCCCGGTGG 1080
      |||
Db      601 CCAACTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGGTGGCATCCCGGTGG 660
QY      1081 CTGATGAGGAGGTGAGAGGCTTGGCTACTCGGTGACTGCAGCCAGCAGGAAATCTTCC 1140
      |||
Db      661 CTGATGAGGAGGTGAGAGGCTTGGCTACTCGGTGACTGCAGCCAGCAGGAAATCTTCC 720
QY      1141 AGGTGGTGAATTCAGAAATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCA 1200
      |||
Db      721 AGGTGGTGAATTCAGAAATCTGCATGGCGATTCCTGCCATGGCCATCCCACTGATCA 780
QY      1201 TGGACACTCTGAGAGAAGAACTTCC 1227
      |||
Db      781 TGGACACTCTGAGAGAAGAACTTCC 807
```

```
RESULT 7
BX334089      1201 bp      mRNA      linear      EST 02-MAY-2003
LOCUS      BX334089 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION  CDNA clone CS0DD006YP08 5-PRIME, mRNA sequence.
ACCESSION  BX334089
VERSION    BX334089.1 GI:30341296
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
```

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5549.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD006DH04QPL&cluster=5549.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DD006DH04QPL.

FEATURES
source
1..1201
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD006YP08"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 266 a 316 c 345 g 245 t 29 others

Query Match 64.6%; Score 795.4; DB 13; Length 1201;
Best Local Similarity 99.3%; Pred. No. 3.7e-186;
Matches 807; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

```
QY      420 GGAAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGAATTCCTTTA 479
      |||
Db      58 GGATGCGGTTCTGAGAGCGTCAGAGAGCGATGGAAGCAAAATGGTGAATTCCTTTA 117
QY      480 GACATCAACATCCAGGAACCTCGCTGGGACCAGAAATCTTTCCTGGGACAGCCCGGCAC 539
      |||
Db      118 GACATCAACATCCAGGAACCTCGCTGGGACCAGAAATCTTTCCTGGGACAGAG-CCGGCAC 176
QY      540 TTTTTCACGTGTACTGATCCTCGAAATCTGCTGCTGTCCGGGGACACAGCTGGAAGCTTCT 599
      |||
Db      177 TTTTTCACGTGTACTGATCCTCGARATCTGCTGCTGTCCGGGGACACAGCTGGAAGCTTCT 236
QY      600 CGGAACATCGTCAGAACTACAGGGCGCGCTGGTACCCCAAGGATCACCCAGAGACCAG 659
      |||
Db      237 CGGAACATCGTCAGAACTACAGGGCGCGCTGGTACCCCAAGGATCACCCAGAGACCAG 296
QY      660 CTGTGAGGGCCCAAGTATGTATGACTCCCGCTTCCATCCGGACACAGGGGAGAGAGGTG 719
      |||
Db      297 CTGTGAGGGCCCAAGTATGTATGACTCCCGCTTCCATCCGGACACAGGGGAGAGAGGTG 356
QY      720 GTCTGATTTGGCCGCGCATGTCAAGCCAGGTGCCATGAACATGACCATCTGCTGATG 779
      |||
Db      357 GTCTGATTTGGCCGCGCATGTCTAGCCAGGTGCCATGAACATGACCATCTGCTGATG 416
QY      780 CTCACATTTCTACAGAGAAGCCCAACCGTGTCTGAGTGGTGAATGATCCTTTC 839
      |||
Db      417 CTCACATTTCTACAGAGAAGCCCAACCGTGTCTGAGTGGTGAATGATCCTTTC 476
QY      840 AATGCCATTTGTTAACTACTCCAAACCGCAGTGTGTACACTCCCATCACTGTAGAGCAGCTG 899
      |||
Db      477 AATGCCATTTGTTAACTACTCCAAACCGCAGTGTGTACACTCCCATCACTGTAGAGCAGCAG 536
QY      900 GGGACAGCCTATGTAGTGGCCACCACCTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCC 959
      |||
Db      537 GGGACAGCCTATGTAGTGGCCACCACCTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCC 596
QY      960 CTCACCAAGCACCCTGCCCCCTTGGTGGCAGATTGTGCCCTTTGCAGCAGTGGCAGCT 1019
      |||
Db      597 CTCACCAAGCACCCTGCCCCCTTGGTGGCAGATTGTGTGCCCTTTGCAGCAGTGGCAGCT 656
QY      1020 GCCAAGTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGGTGGGCATCCCGGTG 1079
      |||
Db      657 GCCAAGTGCATCAACATCCCCCTGATGAGGAGAGAGAGCTGCAGGTGGGCATCCCGGTG 716
QY      1080 GCTGATGAGGAGGTGAGAGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
      |||
Db      717 GCTGATGAGGAGGTGAGAGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
QY      1140 CAGGTGGTGAATTCAGAAATCTGCATGGCGATTCTTCCCATGGCCATCCCACTGATC 1199
      |||
Db      777 CAGGTGGTGAATTCAGAAATCTGCATGGCGATTCTTCCCATGGCCATCCCACTGATC 836
QY      1200 ATGACACTCTGAGAGAAGAAAGACTTCTGAAG 1232
      |||
Db      837 ATGACACTCTGAGAGAAGAAAGACTTCTGAAG 869
```

RESULT 8
BI908654
LOCUS BI908654 777 bp mRNA linear EST 16-OCT-2001
DEFINITION 603066774F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5215818 5',
mRNA sequence.
ACCESSION BI908654
VERSION BI908654.1 GI:16171675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: csabps-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM1542 row: c column: 19
 High quality sequence stop: 771.

FEATURES

source
 1. .777
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5215818"
 /tissue_type="leukocyte"
 /lab_host="DH108"
 /clone_lib="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."

BASE COUNT 141 a 233 c 247 g 156 t
 ORIGIN

Query Match 62.8%; Score 773.8; DB 12; Length 777;
 Best Local Similarity 99.7%; Pred. No. 6.9e-181;
 Matches 775; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 14 GGGACAGCCGAGCGTTACCTGGTCCGGGAGCGAGTTCTTTACCCACCCAGTTCTGG 73
Db 1 GGGACAGCCGAGCGTTACCTGGTCCGGGAGCGAGTTCTTTACCCACCCAGTTCTGG 60
QY 74 TTCTGACGCCCTAGCTCATTCGCAATTAGGGCTTGGGCTTGCTGTCCCTCCG 133
Db 61 TTCTGACGCCCTAGCTCATTCGCAATTAGGGCTTGGGCTTGCTGTCCCTCCG 120
QY 134 CTGGAACCACTCTTCTGTAGCCGAGCCAGCTACCGGGGCTCTGGAATGCCACCCCT 193
Db 121 CTGGAACCACTCTTCTGTAGCCGAGCCAGCTACCGGGGCTCTGGAATGCCACCCCT 180
QY 194 CCGTGGCACCCCTTGAGGCCCTCCGTGAGGGAGACGTACACGGGGGAGAGCGGAGCTGAGCC 253
Db 181 CCGTGGCACCCCTTGAGGCCCTCCGTGAGGGAGACGTACACGGGGGAGAGCGGAGCTGAGCC 240
QY 254 TGAGTTTGCTGACAGCGCTGTGTGTGTGGTGGTCTGCTGCAATCCCGTGCCACC 313
Db 241 TGAGTTTGCTGACAGCGCTGTGTGTGTGGTGGTCTGCTGCAATCCCGTGCCACC 300
QY 314 GGGTGGGCGCGCGGGAAGCTCTGCCCTCCCTGCTGTGGCGCTCACGCGTGAGCTC 373
Db 301 GGGTGGGCGCGCGGGAAGCTCTGCCCTCCCTGCTGTGGCGCTCACGCGTGAGCTC 360
QY 374 CCGCTGATGCTGGAGGGCGCGCGCGAGACAGCGAGGAGAGAGAGAGCGGTTCTG 433
Db 361 CCGCTGATGCTGGAGGGCGCGCGCGAGACAGCGAGGAGAGAGAGAGCGGTTCTG 420
QY 434 AGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGGAATGCCCTTTAGACATCAACATCCA 493
Db 421 AGAGCTTCAGAGAGCGATGGAAGCAAAATGGTGGAATGCCCTTTAGACATCAACATCCA 480
QY 494 GGAACCTCGCTGGGACCAAGTACTTTCTGTGGGAGAGCCCGGCACTTTTCACTGTAC 553
Db 481 GGAACCTCGCTGGGACCAAGTACTTTCTGTGGGAGAGCCCGGCACTTTTCACTGTAC 540
QY 554 TGATCTTCGAATCTGCTGCTGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGCA 613

```

```

Db 541 TGATCTTCGAATCTGCTGCTGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGCA 600
QY 614 GAACCTACAGGGCGCGCGTGTGTGACCCAGGATCACCAGAGACCACTGTGAGGGCCAA 673
Db 601 GAACCTACAGGGCGCGCGTGTGTGACCCAGGATCACCAGAGACCACTGTGAGGGCCAA 660
QY 674 GATGTGTATGACTCCGCTTCCATCCGACACAGGGGAGAGAGTGTCTGTATGGCCG 733
Db 661 GATGTGTATGACTCCGACTTCCATCCGACACAGGGGAGAGAGTGTCTGTATGGCCG 720
QY 734 CATGTACAGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATCTTA 790
Db 721 CATGTACAGCCAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATCTTA 777

```

RESULT 9
 AL522228 840 bp mRNA linear EST 22-MAY-2003
 LOCUS AL522228
 DEFINITION AL522228 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
 CDNA clone CS0DB007YP05 5-PRIME, mRNA sequence.
 ACCESSION AL522228
 VERSION AL522228.2 GI:31040500
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12785721.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f For
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DB007CH03QPI&cluster=5549.f)
 feng liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue Genoscope sequence ID : CS0DB007CH03QPI.

FEATURES

source
 1. .840
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB007YP05"
 /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 162 a 245 c 263 g 170 t
 ORIGIN

Query Match 62.3%; Score 767.2; DB 9; Length 840;
 Best Local Similarity 97.6%; Pred. No. 3.1e-179;
 Matches 789; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

```

QY 6 AGGCGCAGGGAGACAGCGGATACCTGTCCCGGAGCGGAGTTCTTACCAACCCC 65
Db 34 AGCAGGCTGTGTACCGGTCCGGAATCCCGGGATCGGGCAGCGGAGTTCTTACCAACCCC 93
QY 66 AGTCTGCTTCTGACGCCCTAGCTCATTCGCAAAATTAGGGCTTGGGCTGTGCTTTC 125
Db 94 AGTCTGCTTCTGACGCCCTAGCTCATTCGCAAAATTAGGGCTTGGGCTGTGCTTTC 153
QY 126 CCGTCCGCTCGAACCACTCTTCTGTAGCCGAGCCAGCTACCGGGGCTCCTGGAATTG 185

```


Db	154	CCCTCCGGCTCGAACCACTCTTCTGTGAGCCGAGCCAGCTACCGGGGCTCTGGAATTG	213
QY	186	CCACCCCTCCCTGGGACCCCTTGAGGCTCCGTGGAGGAGCTCACGGGGCAGACGGGA	245
Db	214	CCACCCCTCCCTGGGACCCCTTGAGGCTCCGTGGAGGAGCTCACGGGGCAGACGGGA	273
QY	246	CGTAGCCTGAGTTTGGCTGCAGGCGTCTCTGTGTGGTGGCTGGGTTTGCCAATCCCCG	305
Db	274	CGTAGCCTGAGTTTGGCTGCAGGCGTCTCTGTGTGGTGGCTGGGTTTGCCAATCCCCG	333
QY	306	TGCCACCCGGGTGGGCGCGGGGAACTCTCTGCCCTCCCTGCTGGTGGCGTCCACGC	365
Db	334	TGCCACCCGGGTGGGCGCGGGGAACTCTCTGCCCTCCCTGCTGGTGGCGTCCACGC	393
QY	366	GTGACGTCCTCCGCTGATGGCTGGAGGGCCCGGGCGGACAGCGGAGCAGAGGAAGG	425
Db	394	GTGACGTCCTCCGCTGATGGCTGGAGGGCCCGGGCGGACAGCGGAGCAGAGGAAGG	453
QY	426	CGGTTCTGAGAGCTTCAGAGAGCGATGGAAGAACAATGGGTGAATTGCCCTTAGACATC	485
Db	454	CGGTTCTGAGAGCTTCAGAGAGCGATGGAAGAACAATGGGTGAATTGCCCTTAGACATC	513
QY	486	AACATCCAGGAACCTCGCTGGGACCAAGTACTTCTCTGGGACAGCCCGCACTTTTTC	545
Db	514	AACATCCAGGAACCTCGCTGGGACCAAGTACTTCTCTGGGACAGCCCGCACTTTTTC	572
QY	546	ACTGTTACTGATCCTCGAAATCTGCTGCTGTCCGGGACACAGCTGGAAGCTTCTCGAAC	605
Db	573	ACTGTTACTGATCCTCGAAATCTGCTGCTGTCCGGGACACAGCTGGAAGCTTCTCGAAC	632
QY	606	ATCGTGCAGAACTACAGGGCGCGGTGTGTGACCCAGGGATCAACGAGACAGCTGTGG	665
Db	633	ATCGTGCAGAACTACAGGGCGCGGTGTGTGACCCAGGGATCAACGAGACAGCTGTGG	692
QY	666	AGGGCCAAGTATGTATGATCTCCGCTTCCATCCGGACACAGGGGAGAAGTGTCTTG	725
Db	693	AGGGCCAAGTATGTATGATCTCCGCTTCCATCCGGACACAGGGGAGAAGTGTCTTG	752
QY	726	ATTGCCCGCATGTACAGCCAGGTGCCCATGACATGACATCATCTGGCTGCATGCTCACA	785
Db	753	ATTGCCCGCATGTACAGCCAGGTGCCCATGACATGACATCATCTGGCTGCATGCTCACA	812
QY	786	TTCTACAGGAAGACCCCAACCGTGGTGT	813
Db	813	TTCTACAGGAAGACCCCAACCGTGGTGT	840
RESULT 10			
LOCUS	BI818854	811 bp	mrna
DEFINITION	603037323F1 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5178437 5',		
ACCESSION	BI818854		
VERSION	BI818854.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 811)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		

```

Plate: L1AM11444   row: n   column: 06
High quality sequence stop: 784.
Location/Qualifiers
1. . 811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5178437"
/lab_host="DH10B"
/clone_11b="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      180 a      223 c      246 g      162 t
ORIGIN

```

Query Match	Best Local Similarity	Score	DB 12:	Length	811:
99.0%	99.0%	759.6	DB 12:	Length	811:
Matches	806;	Conservative	0;	Mismatches	4;
				Indels	4;
				Gaps	4;
QY	378	GTGATGCGCTGGAGGGCCCGGGCGACAGCGGAGGACAGAGAGGAAGCGCGTTCGAGAG	437		
Db	1	GTGATGCGCTGGAGGGCCCGGGCGACAGCGGAGGACAGAGAGGAAGCGCGTTCGAGAG	60		
QY	438	CTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGA	497		
Db	61	CTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGA	119		
QY	498	CCTCGCTGGGACCAAGTACTTTCTGGGACAGAGCCCGCACTTTTCACTGTTACTGAT	557		
Db	120	CCTCGCTGGGACCAAGTACTTTCTGGGACAGAGCCCGCACTTTTCACTGTTACTGAT	178		
QY	558	CCTCGAATCTGCTGCTGCTCCGGGGGCACAGCTGGAAGCTTCTCGGAACATCGTCAGAAC	617		
Db	179	CCTCGAATCTGCTGCTGCTCCGGGGGCACAGCTGGAAGCTTCTCGGAACATCGTCAGAAC	238		
QY	618	TACAGGGCCGGCGTGGTGACCCCGAGGATCACCGAGACCAGCTGTGGAGGCCAAGTAT	677		
Db	239	TACAGGGCCGGCGTGGTGACCCCGAGGATCACCGAGACCAGCTGTGGAGGCCAAGTAT	298		
QY	678	GTGTATGACTCCGCTTCCATCCGACACAGGGGAGAAAGTGTCTGATTTGGCCGATG	737		
Db	299	GTGTATGACTCCGCTTCCATCCGACACAGGGGAGAAAGTGTCTGATTTGGCCGATG	358		
QY	738	TCAGCCGAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAG	797		
Db	359	TCAGCCGAGGTGCCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAG	418		
QY	798	ACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCCCTCAATGCCATTGTTAACTAC	857		
Db	419	ACCCCAACCGTGTGTTCTGGCAGTGGGTGAATCAGTCCCTCAATGCCATTGTTAACTAC	478		
QY	858	TCCAACCGCAGTGTGTACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTGAGT	917		
Db	479	TCCAACCGCAGTGTGTACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTGAGT	538		
QY	918	GCCACCACTGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTACACCAAGCACTGCCCC	977		
Db	539	GCCACCACTGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTACACCAAGCACTGCCCC	598		
QY	978	CCCTGTGTCGGCAGATTGTGCCCTTTGACAGAGTGGCAGCTGCCAATGCATCAACATC	1037		
Db	599	CCCTGTGTCGGCAGATTGTGCCCTTTGACAGAGTGGCAGCTGCCAATGCATCAACATC	658		
QY	1038	CCCCGTATGAGGCAGAGAGCTGCAGTGGGCATCCCGGTGGCTATGAGGCAGTCAAG	1097		


```

Db      659  CCCCTGATGAGCAGAGAGAGCTGCAGTGGGCATCCCGGTGGCTGATGAGCAGTTCAG 718
QY      1098 AGCCTTGCTACTCGGTGACTGCAGCCAG-CAGGAATCTTCCAGGTGTGATTTCAAG 1156
        |||||||
Db      719  AGGCTTGCTACTCGGTGACTGCAGACAGCAGGAATCTTCCAGGTGTGATTTCAAG 778
QY      1157 AATCTGCATGGCGATTCCTGCATGGCCATCCCA 1190
        |||||||
Db      779  AATCTGCATGGCGA-TCTGCATGGGCATCCCA 811

RESULT 11
BM925231      1076 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT_6627462 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5762556
DEFINITION  5', mRNA sequence.
ACCESSION  BM925231
VERSION    BM925231.1 GI:19375610
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1076)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12812 row: P column: 13
            High quality sequence start: 4
            High quality sequence stop: 664.

FEATURES
    source
        1..1076
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5762556"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_122"
            /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
            Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
            anonymous pool of 24 week female lung, 16 week female
            spleen, and 20-22 week male spleens. Library is oligo-dT
            primed and directionally cloned (EcoRV site is destroyed
            upon cloning). Average insert size 1.4 kb, insert size
            range 1-3 kb. Library is normalized and enriched for
            full-length clones and was constructed by C. Gruber
            (Invitrogen). Research Genetics tracking code 026. Note:
            this is a NIH-MGC Library."

BASE COUNT      230 a      316 c      314 g      216 t
ORIGIN
Query Match      61.6%; Score 759.4; DB 12; Length 1076;
Best Local Similarity 99.7%; Pred. No. 2.9e-177;
Matches 771; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      460  AATGGGTGAATGCTTACATCAACATCCAGGAACCTGCTGGACCAAGTACTT 519
        |||||||
Db      11  AATGGGTGAATGCTTACATC-ACATCCAGGAACCTGCTGGACCAAGTACTT 69
QY      520  TCCCTGGGCGAGAGCCCGGACATTTTCACTGTACTGATGATCTCGAAATCTGCTGCTCCG 579
        |||||||
Db      70  TCCCTGGGCGAGAGCCCGGACATTTTCACTGTACTGATGATCTCGAAATCTGCTGCTCCG 129
QY      580  GGGCAGAGCTGGAAGCTTCTCGGAACATCGTGAGAACTACAGGCGCGGTGTGACCC 639
```

```

Db      130  GGGCAGAGCTGGAAGCTTCTCGGAACATCGTGAGAACTACAGGCGCGGTGTGACCC 189
QY      640  CAGGATCACCGAGGACAGCTGTGAGGGGCCAAGTATGTATGACTCCGCTTCATC 699
        |||||||
Db      190  CAGGATCACCGAGGACAGCTGTGAGGGGCCAAGTATGTATGACTCCGCTTCATC 249
QY      700  CGGACACAGGGGAGAAGGTGCTCTGATTGGCCGATGTCAGCCAGGTGCCATGAACA 759
        |||||||
Db      250  CGGACACAGGGGAGAAGGTGCTCTGATTGGCCGATGTCAGCCAGGTGCCATGAACA 309
QY      760  TGACCATCAGCTGGCTGCATGCTCACAATTCTACAGGAAGACCCCAACCGTGTCTGCG 819
        |||||||
Db      310  TGACCATCAGCTGGCTGCATGCTCACAATTCTACAGGAAGACCCCAACCGTGTCTGCG 369
QY      820  AGTGGGTGAATCAGTCTCTCAATGCCATTGTTACTACTCCAACCCGAGTGTGACACTC 879
        |||||||
Db      370  AGTGGGTGAATCAGTCTCTCAATGCCATTGTTACTACTCCAACCCGAGTGTGACACTC 429
QY      880  CCATCACTGTGAGGAGCTGGGGACAGCCTATGTAGTGGCCACACTGGAGCTGTGGCCA 939
        |||||||
Db      430  CCATCACTGTGAGGAGCTGGGGACAGCCTATGTAGTGGCCACACTGGAGCTGTGGCCA 489
QY      940  CGGCCCTGGGACTCAATCCCTCACCAGACACCTGCCCCCTTGGTGGCAGATTGTGTC 999
        |||||||
Db      490  CGGCCCTGGGACTCAATCCCTCACCAGACACCTGCCCCCTTGGTGGCAGATTGTGTC 549
QY      1000 CTTTGCAGCAGTGGCAGCTGCCAAGCTGCATCAACATCCCCCTGATGAGGACAGAGAGAGC 1059
        |||||||
Db      550  CTTTGCAGCAGTGGCAGCTGCCAAGCTGCATCAACATCCCCCTGATGAGGACAGAGAGAGC 609
QY      1060 TGCAGTGGGGCATCCCGGTGGCTGATGAGGACAGTCCAGAGGCTTGGCTACTCGGTGACTG 1119
        |||||||
Db      610  TGCAGTGGGGCATCCCGGTGGCTGATGAGGACAGTCCAGAGGCTTGGCTACTCGGTGACTG 669
QY      1120 CAGCCACGACGAGGAATCTTCCAGGTGCTGATTTCAGAAGATCTGCATGGCGGATTCCTGCCA 1179
        |||||||
Db      670  CAGCCACGACGAGGAATCTTCCAGGTGCTGATTTCAGAAGATCTGCATGGCGGATTCCTGCCA 729
QY      1180 TGGCCATCCCAACCACTGATCATGACACTCTGGAGAAGAAAGACTTCTTGAAG 1232
        |||||||
Db      730  TGGCCATCCCAACCACTGATCATGACACTCTGGAGAAGAAAGACTTCTTGAAG 782

RESULT 12
BM917821      1014 bp      mRNA      linear      EST 16-JUL-2002
LOCUS      AGENCOURT_8229525 lupski_dorsal_root_ganglion Homo sapiens cDNA
DEFINITION  clone IMAGE:6184833 5', mRNA sequence.
ACCESSION  BM917821
VERSION    BM917821.1 GI:21856718
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1014)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13574 row: K column: 10
            High quality sequence stop: 610.

FEATURES
    source
        1..1014
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184833"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTTAGATCGCAGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT      199 a      292 c      317 g      206 t
ORIGIN
Query Match      59.9%; Score 738.4; DB 13; Length 1014;
Best Local Similarity 96.5%; Pred. No. 4.5e-172;
Matches 799; Conservative 0; Mismatches 21; Indels 8; Gaps 4;
QY 222 GGGACGTCACGGGGCAGAGCGGAGCGTGAAGCTTGTGTCAGAGCGTCTGTGTG 281
Db 22 GGGATGTACGGGGCAGAGCGGAGCGTGAAGCTTGTGTCAGAGCGTCTGTGTG 81
QY 282 GTGGCTGGGTCTGCCAATCCCGTGGCCACCGGGTGGCGCGCGCGGGAAGCTCTGCC 341
Db 82 GTGGCTGGGTCTGCCAATCCCGTGGCCACCGGGTGGCGCGCGCGGGAAGCTCTGCC 141
QY 342 CCTCCCTGCTGTCGGCGTCAAGCGTCCCGGTGATGGCTGGAGGGCCCGCGG 401
Db 142 CCTCCCTGCTGTCGGCGTCAAGCGTCCCGGTGATGGCTGGAGGGCCCGCGG 201
QY 402 CGACAGCGGAGGAGAGAGAGAGGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAA 461
Db 202 CGACAGCGGAGGAGAGAGAGAGGCGGTCTGAGAGCTTCAGAGAGCGATGGAAGCAA 261
QY 462 ATGGTGAATTGCTTTAGACATCAACATCCAGAACCTCGCTGGAGCCAAAGTACTTTC 521
Db 262 ATGGTGAATTGCTTTAGACATCAACATCCAGAACCTCGCTGGAGCCAAAGTACTTTC 321
QY 522 CTGGGAGAGCCCGGCACTTTTCTACTGTACTGATCTCGAAATCTGCTGTCTCGGG 581
Db 322 CTGGGAGAGCCCGGCACTTTTCTACTGTACTGATCTCGAAATCTGCTGTCTCGGG 381
QY 582 GCACAGCTGGAAGCTTCTCGGAACATGTCAGAACTACAGGGCCGCGTGTGACCCCA 641
Db 382 GCACAGCTGGAAGCTTCTCGGAACATGTCAGAACTACAGGGCCGCGTGTGACCCCA 441
QY 642 GGGATCACCGAGGAGCAGCTGTGGAGGGCCAAAGTATGTATGACTCCGCTTCATCCG 701
Db 442 GGGATCACCGAGGAGCAGCTGTGGAGGGCCAAAGTATGTATGACTCCGCTTCATCCG 501
QY 702 GACACAGGGGAGAGGTTGCTGTATGGCCGATGTACGCCAGGTGCCCATGAACATG 761
Db 502 GACACAGGGGAGAGGTTGCTGTATGGCCGATGTACGCCAGGTGCCCATGAACATG 561
QY 762 ACCATCACTGGCTGATGTACATTTCTACAGGAAGAGCCCAACCGTGTGTCTGGCAG 821
Db 562 ACCATCACTGGCTGATGTACATTTCTACAGGAAGAGCCCAACCGTGTGTCTGGCAG 621
QY 822 TGGGTGAATCAGTCTTCAATGCCATTTGTTACTACTTCAACCGCAGTGGTGACACTCC 881
Db 622 TGGGTGAATCAGTCTTCAATGCCATTTGTTACTACTTCAACCGCAGTGGTGACACTCC 681
QY 882 ATCACTGTGAGGAGCTGGGAGACCCCTATGTAGTGCACCACTGGAGCTGTGGCAGG 941
Db 682 ATCACTGTGAGGAGCTGGGAGACCCCTATGTAGTGCACCACTGGAGCTGTGGCAGG 741
```

```
QY 942 GCCCTGGAGCTC-AAATCCTTCAACCAAGCACCCTGCCCC---TTGGTCGGCAGATT--- 994
Db 742 GCCCTGGAGCTCAAAATCCTTCAACCAAGCACCCTGCCCCCTTTGGTCGGCAGATTGG 801
QY 995 TGTGCCCTTTGACAGCAGTGGCAGCTG-CCAAGTGCATCAACATCCCC 1041
Db 802 TGCCCTTTGACAGCAGTGGCAGCTGTCAAACTGCATTCACATCCCC 849

RESULT 13
BQ425697
LOCUS
DEFINITION
BQ425697 862 bp mRNA linear EST 23-MAY-2002
AGENCOURT_7887371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6166944
5', mRNA sequence.
ACCESSION
BQ425697
VERSION
BQ425697.1 GI:21121012
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 862)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AMJ3528 row: b column: 01
High quality sequence stop: 603.
FEATURES
source
1..862
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6166944"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT      200 a      249 c      246 g      166 t      1 others
ORIGIN
Query Match      59.0%; Score 727.2; DB 13; Length 862;
Best Local Similarity 98.7%; Pred. No. 2.5e-169;
Matches 743; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 395 CCGCGCGGACAGCGGAGGAGAGAGAGGCGGTTCTGAGAGCTTCAGAGCGATGGA 454
Db 1 CCGCGCGGACAGCGGAGGAGAGAGAGGCGGTTCTGAGAGCTTCAGAGCGATGGA 60
QY 455 AAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGAGCCAAAG 514
Db 61 AAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGAGCCAAAG 120
QY 515 TACTTTCCTGGGAGAGCGCGGCACTTTTCACTGTACTGATCTCGAAATCTGCTGCT 574
Db 121 TACTTTCCTGGGAGAGCGCGGCACTTTTCACTGTACTGATCTCGAAATCTGCTGCT 180
QY 575 GTCCGGGACACAGCTGGAAGCTTCTCGGAACATGTCGAGAACTACAGAGGCGCGGTGT 634
Db 181 GTCCGGGACACAGCTGGAAGCTTCTCGGAACATGTCGAGAACTACAGAGGCGCGGTGT 240
```

```
QY 635 GACCCAGGATCACCGAGGACAGCTGTGGAGGGCCAGATATGTATGACTCCGCTT 694
    |||||||
Db 241 GACCCAGGATCACCGAGGACAGCTGTGGAGGGCCAGATATGTATGACTCCGCTT 300
QY 695 CCATCCGGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTCAGCCAGTGCCCAT 754
    |||||||
Db 301 CCATCCGGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTCAGCCAGTGCCCAT 360
QY 755 GAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGCTT 814
    |||||||
Db 361 GAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGCTT 420
QY 815 CTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAACTACTCCACCGAGTGTA 874
    |||||||
Db 421 CTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAACTACTCCACCGAGTGTA 480
QY 875 CACTCCCATCACTGTGAGGACAGCTGGGACAGCCCTATGTAGTGCCA-CCACTGAGCTG 933
    |||||||
Db 481 CACTCCCATCACTGTGAGGACAGCTGGGACAGCCCTATGTAGTGCCA-CCACTGAGCTG 540
QY 934 TGGCCAGGCCCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGCTGGCAGAT 993
    |||||||
Db 541 TGGCCAGGCCCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGCTGGCAGAT 600
QY 994 TTGTGCCCCCTTGACAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGACA 1053
    |||||||
Db 601 TTGTGCCCCCTTGACAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGACA 660
QY 1054 GAGAGCTGCAGGTGGGATCCCCGGTGGCTGATGAGGAGGTCAGAGGCTTGCTACTCGG 1113
    |||||||
Db 661 GAGAGCTGCANGTGGGATCCCCGGTGGCTGATGAGGAGGTCAGAGGCTTGCTACTCGG 720
QY 1114 TGACTGCAGCCCAAGCAGGGAATCTTCCAGGTGG 1146
    |||||||
Db 721 TGACTGCAGCCCAAGCAGGGAATCTTCCAGGTGG 753

RESULT 14
BO959114 963 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_10031240 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482227
5', mRNA sequence.
BO959114
ACCESSION BO959114.1 GI:22374592
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 963)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2664 row: n column: 20
High quality sequence stop: 635.
Location/Qualifiers
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6482227"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
```

```
BASE COUNT 200 a 290 c 272 g 201 t
ORIGIN
Query Match 57.9%; Score 713; DB 13; Length 963;
Best Local Similarity 100.0%; Pred. No. 8.6e-166;
Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TCCTGGGACAGCCCCGGCACTTTTCACTGTTACTGATCCTCGAAATCTGCTGTCGG 579
    |||||||
Db 1 TCCTGGGACAGCCCCGGCACTTTTCACTGTTACTGATCCTCGAAATCTGCTGTCGG 60
QY 580 GGGCACAGCTGGAAGCTTCTCGGAACATCTGCAGAACTACAGGGCCGGCGTGTCACCC 639
    |||||||
Db 61 GGGCACAGCTGGAAGCTTCTCGGAACATCTGCAGAACTACAGGGCCGGCGTGTCACCC 120
QY 640 CAGGATCACCGAGACAGCTGTGGAGGGCCAAAGTATGTATGATGATCTCCGCTTCCATC 699
    |||||||
Db 121 CAGGATCACCGAGACAGCTGTGGAGGGCCAAAGTATGTATGATGATCTCCGCTTCCATC 180
QY 700 CGGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTACAGCCAGTGGCCATGAACA 759
    |||||||
Db 181 CGGACACAGGGGAGAGAGTGTCTGATTTGGCCGATGTACAGCCAGTGGCCATGAACA 240
QY 760 TGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTCTGGC 819
    |||||||
Db 241 TGACCATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTCTGGC 300
QY 820 AGTGGGTGAATCAGTCTTCAATGCCATTGTTAACTACTCCAAACCGAGTGTGACACTC 879
    |||||||
Db 301 AGTGGGTGAATCAGTCTTCAATGCCATTGTTAACTACTCCAAACCGAGTGTGACACTC 360
QY 880 CCATCACTGTAGGACAGCTGGGACAGCCTATGTAGTGCCACCACTGAGCTGTGGCCA 939
    |||||||
Db 361 CCATCACTGTAGGACAGCTGGGACAGCCTATGTAGTGCCACCACTGAGCTGTGGCCA 420
QY 940 CGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGCTGGCAGATTGTGC 999
    |||||||
Db 421 CGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGCTGGCAGATTGTGC 480
QY 1000 CCTTTCAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGACAGAGAGC 1059
    |||||||
Db 481 CCTTTCAGCAGTGGCAGCTGCCAATCTGCATCAACATCCCCCTGATGAGGACAGAGAGC 540
QY 1060 TGCAGGTGGGCATCCCGGTGGCTGATGAGGCGAGGTCAGAGGCTTGCTACTCGGTGACTG 1119
    |||||||
Db 541 TGCAGGTGGGCATCCCGGTGGCTGATGAGGCGAGGTCAGAGGCTTGCTACTCGGTGACTG 600
QY 1120 CAGCCAAAGCAGGAATCTTCCAGGTGTGATTTCAAGAAATCTGATGGCATTCCTGCCA 1179
    |||||||
Db 601 CAGCCAAAGCAGGAATCTTCCAGGTGTGATTTCAAGAAATCTGATGGCATTCCTGCCA 660
QY 1180 TGGCCATCCCAACCACTGATCATGACACTCTGGAGAAAGAACTTCTGGAAG 1232
    |||||||
Db 661 TGGCCATCCCAACCACTGATCATGACACTCTGGAGAAAGAACTTCTGGAAG 713

RESULT 15
BX393485 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX393485
DEFINITION BX393485 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
CDNA clone CS0DB004YM02 5-PRIME, mRNA sequence.
BX393485
ACCESSION BX393485
VERSION BX393485.1 GI:30616089
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f. For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DB004BG01QPI&cluster=5549.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DB004BG01QPI.
 Location/Qualifiers
 1..1201

FEATURES
 source ..
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB004YM02"
 /issue_type="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone_1ib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 48 a 384 c 428 g 276 t 65 others
 ORIGIN

Query Match 57.8%; Score 711.6; DB 13; Length 1201;
 Best Local Similarity 83.9%; Pred. No. 2.1e-165;
 Matches 794; Conservative 15; Mismatches 136; Indels 1; Gaps 1;

QY 15 GGACAGCCGAGCGTTACCTGCTCCCGGAGGAGTCTTACCCACCCAGTTCTGT 74
 Db 57 GGGGTGCGGGCGTNCCTGCTCCCGGCGGGGCTTCTTCCCGCCCGCTTGTGT 116
 QY 75 TCTGACGCCCTAGCTCATTCGCAATTAGGCTTGGCTGCTGCTTCCCTCCGCG 134
 Db 117 TCTGCGCCCTGCTCTTCCGCCCTTTCGGGCTTGGCTGCTGCTTCCCTCCGCG 176
 QY 135 TCGAACCACTCTTCTCTGAGCCGAGCCAGTACCGGGCTCCTGGAATTGCCACCCCTC 194
 Db 177 TCGCCCCCTCTTCTCTGCGCGCGCCCTCCCGGGCTCCTGCTTGCCTCCCGCTC 236
 QY 195 CCTGGGCACTTGAGGCTCCGTGAGGAGCTCACGGGGGAGAGAGGAGCGTGAAGCT 254
 Db 237 CCTGGGCCCCCTTGGGCTCCGTGAGGAGCTCCCGGGGAGGAGGAGGAGGAGGCT 296
 QY 255 GAGTTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
 Db 297 GGGTTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 356
 QY 315 GGTGGGCGGCGGAGAGCTCCTGCTCCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCT 374
 Db 357 GGTGGGCGGCGGCGGCGGCTCCTGCTCCTCCCTCCCTGCTGCTGCTGCTGCTGCT 416
 QY 375 CGCGTATGCTGAGGAGGCGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 434
 Db 417 CGCGTATGCTGAGGAGGCGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 476
 QY 435 GAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTGCTTGAACATCAACATCCAG 494
 Db 477 GGGCTTCGGGGGGGGGCTTGGGGGGGGGCTTGGGGGGGCTTGGGGGGGCTTGGGG 536
 QY 495 GAACCTGCTGGGAGCAAAAGTACTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 554
 Db 537 GCCCTTCGCTGGGCGCGCGCTTCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 596

QY 555 GATCCTCGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
 Db 597 GTTCCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
 QY 615 AACTACAGGGCGGCGGTGAGACCCAGGATCACCGAGACCAAGCTGTGAGGGCCAG 674
 Db 657 AACTACAGGGCGGCGGTGAGACCCAGGATCACCGAGACCAAGCTGTGAGGGCCAG 716
 QY 675 TATGTATGACTCCGCTTCCATCCGAGACACAGGGAGAGAGTGTCTGATGTGCCCC 734
 Db 717 TTGTGTGTGACTCCGCTTCCATCCGAGACACAGGGAGAGTGTCTGATGTGCCCC 776
 QY 735 ATGTACAGCCAGGTGCCATGAACATGACCATCATGCTGCTGCTGCTGCTGCTGCTGCT 794
 Db 777 CTGTACAGCCAGGTGCCATGAACATGACCATCATGCTGCTGCTGCTGCTGCTGCTGCT 836
 QY 795 AAGACCCCAACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
 Db 837 RGGGCCCCACCGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
 QY 855 TACTCCACCCAGTGTGACACTCCATCATCTGAGGAGAGCTGGGAGACAGCTATGTG 914
 Db 897 TTCTCCMCCCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
 QY 915 AGTCCACCACTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 Db 957 KGTGCC-CCCCGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1001

Search completed: October 16, 2003, 15:58:02
 Job time : 2511 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 16:06:27 ; Search time 261 Seconds

(without alignments)
2626.128 Million cell updates/sec

Title: US-09-990-415A-2

Perfect score: 1345
Sequence: 1 MESKMGEPLDINIQEPRWD.....PAMAIPLIMDTLEKKDPLK 261

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Egapop 6.0	Egapext 7.0
	Delop 6.0	Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US09990415/runat_16102003_073808_9492/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09990415_ecgn1_1_85_arunat_16102003_073808_9492
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1232	10	US-09-990-415A-1 Sequence 1, Appl1

2	1345	100.0	1375	9	US-09-867-550-1343	Sequence 1343, Ap
3	1345	100.0	1729	14	US-10-037-270-1016	Sequence 1016, Ap
4	1089	81.0	1134	14	US-10-014-338-1	Sequence 1, Appl1
5	1089	81.0	2269	10	US-09-990-415A-7	Sequence 7, Appl1
6	1089	81.0	2978	14	US-10-198-846-9767	Sequence 9767, Ap
7	1003.5	74.6	1061	10	US-09-990-415A-3	Sequence 3, Appl1
8	897.5	66.7	984	14	US-10-014-338-3	Sequence 3, Appl1
9	754	56.1	1567	10	US-09-990-415A-5	Sequence 5, Appl1
10	750	55.8	2376	13	US-10-098-841-64	Sequence 64, Appl1
11	750	55.8	2458	9	US-09-822-849A-490	Sequence 490, App
12	708	52.6	670	14	US-10-206-901B-41	Sequence 41, Appl
13	587	43.6	489	11	US-09-918-995-23576	Sequence 23576, A
14	447.5	33.3	3342	12	US-09-814-353-20153	Sequence 20153, A
15	358	26.6	522	11	US-09-918-995-5675	Sequence 5675, A
16	344	25.6	293	10	US-09-783-590-2110	Sequence 2110, Ap
17	321	23.9	648	14	US-10-198-846-7595	Sequence 7595, Ap
18	295	21.9	533	11	US-09-918-995-26900	Sequence 26900, A
19	230	17.1	238	11	US-09-991-936-1857	Sequence 1857, Ap
20	211	15.7	496	14	US-10-106-698-2663	Sequence 2663, Ap
21	97	7.2	549	11	US-09-991-936-1707	Sequence 1707, Ap
22	96.5	7.2	966	11	US-09-974-879-25	Sequence 25, Appl
23	96.5	7.2	966	11	US-09-305-736-25	Sequence 25, Appl
24	96	7.1	3139	12	US-10-252-157-457	Sequence 457, App
25	88.5	6.6	1061	14	US-10-161-572-19	Sequence 19, Appl
26	88.5	6.6	1094	14	US-10-161-572-17	Sequence 17, Appl
27	88.5	6.6	1130	14	US-10-161-572-18	Sequence 18, Appl
28	88.5	6.6	1603	10	US-09-966-546-3	Sequence 3, Appl1
29	88.5	6.6	1603	10	US-09-966-545-3	Sequence 3, Appl1
30	88.5	6.6	1603	11	US-09-965-212-3	Sequence 3, Appl1
31	88.5	6.6	1603	15	US-10-189-940-3	Sequence 72, Appl
32	88.5	6.6	1678	13	US-10-098-841-72	Sequence 72, Appl
33	88.5	6.6	1679	10	US-09-978-295A-522	Sequence 522, App
34	88.5	6.6	1679	10	US-09-978-697-522	Sequence 522, App
35	88.5	6.6	1679	10	US-09-978-192A-522	Sequence 522, App
36	88.5	6.6	1679	10	US-09-999-832A-522	Sequence 522, App
37	88.5	6.6	1679	11	US-09-978-189-522	Sequence 522, App
38	88.5	6.6	1679	11	US-09-978-608A-522	Sequence 522, App
39	88.5	6.6	1679	11	US-09-978-585A-522	Sequence 522, App
40	88.5	6.6	1679	11	US-09-978-191A-522	Sequence 522, App
41	88.5	6.6	1679	11	US-09-978-403A-522	Sequence 522, App
42	88.5	6.6	1679	11	US-09-978-564A-522	Sequence 522, App
43	88.5	6.6	1679	11	US-09-999-833A-522	Sequence 522, App
44	88.5	6.6	1679	11	US-09-981-915A-522	Sequence 522, App
45	88.5	6.6	1679	11	US-09-978-824-522	Sequence 522, App

ALIGNMENTS

RESULT 1
US-09-990-415A-1
; Sequence 1, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (450)..(1232)
US-09-990-415A-1

Alignment Scores:
Pred. No.: 3.19e-171
Score: 1345.00
Percent Similarity: 100.00%

Length: 1232
Matches: 261
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-990-415A-1 (1-1232)

```
QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIleProArgTyrPasp 20
      |||||||
Db      450 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 509
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
      |||||||
Db      510 CAAAGTACTTCTCTGGGAGAGAGCCCGGACATTTCAGTGTACTGATCCTCGAAATCTG 569
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
      |||||||
Db      570 CTGCTGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGACAGAACTACAGGCGCGGC 629
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80
      |||||||
Db      630 GTGGTGACCCAGGAGATCAACGAGAGACAGCTGTGAGGGCCCAAGTATGTATGACTCC 689
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
      |||||||
Db      690 GCCTTCCATCCGACACAGGGGAGAAGGTGCTCTGATTGGCCGCGCATGTACAGCCGAGTG 749
QY      101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
      |||||||
Db      750 CCCATGAACATGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
QY      121 ValPheTyrPglIntPyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
      |||||||
Db      810 GTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTAACTACTCAACCGCAGT 869
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
      |||||||
Db      870 GGTGACACTCCCATCTACTGTGAGGACAGCTGGGAGACCTATGTAGTGGCCACCACTGGA 929
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
      |||||||
Db      930 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCTTGTGCGGC 989
QY      181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
      |||||||
Db      990 AGATTGTGCCCCCTTGGCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGG 1049
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
      |||||||
Db      1050 CAGAGAGAGCTGCAGTGGGCATCCCGGTGGCTGATGAGGACAGGTGAGAGGCTTGGCTAC 1109
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
      |||||||
Db      1110 TCGGTGACTGCAGCCAGCAGGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCG 1169
QY      241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
      |||||||
Db      1170 ATTCTGCGCATGGCCATCCACCACCTGATCATGAGCACTCTGAGAGAAAGACTTCTCTG 1229
QY      261 Lys 261
      |||
Db      1230 AAG 1232
```

RESULT 2
US-09-867-550-1343

; Sequence 1343, Application US/09867550
; Patent No. US20020082206A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Mehraban, Fuad,

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

```
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1343
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(2)
; OTHER INFORMATION: wherein any n is one of a or t or c or g
US-09-867-550-1343
```

Alignment Scores:

Pred. No.:	3,79e-171	Length:	1375
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-990-415A-2 (1-261) x US-09-867-550-1343 (1-1375)

```
QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIleProArgTyrPasp 20
      |||||||
Db      91 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 150
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
      |||||||
Db      151 CAAAGTACTTCTCTGGGAGAGAGCCCGGACATTTCAGTGTACTGATCCTCGAAATCTG 210
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
      |||||||
Db      211 CTGCTGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGACAGAACTACAGGCGCGGC 270
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80
      |||||||
Db      271 GTGGTGACCCAGGAGATCAACGAGAGACAGCTGTGAGGGCCCAAGTATGTATGACTCC 330
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
      |||||||
Db      331 GCCTTCCATCCGACACAGGGGAGAAGGTGCTCTGATTGGCCGCGCATGTACAGCCGAGTG 390
QY      101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
      |||||||
Db      391 CCCATGAACATGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 450
QY      121 ValPheTyrPglIntPyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
      |||||||
Db      451 GTGTTTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTAACTACTCCAAACCGCAGT 510
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
      |||||||
Db      511 GGTGACACTCCCATCTACTGTGAGGACAGCTGGGAGACGCTATGTAGTGCACCACTGGA 570
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
      |||||||
Db      571 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCTTGTGTCGGC 630
QY      181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
      |||||||
Db      631 AGATTGTGCCCTTGGCAGCAGTGGCAGCTGCCCACTGCATCAACATCCCCCTGATGAGG 690
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
      |||||||
Db      691 CAGAGAGAGCTGCAGTGGGCATCCCGGTGGCTGATGAGGACAGGTGAGAGGCTTGGCTAC 750
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
```

```

; ; FEATURE:
; ; NAME/KEY: misc_feature
; ; LOCATION: (1132)..(1133)
; ; OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-1

```

Alignment Scores:

Pred. No.:	1.06e-136	Length:	1134
Score:	1089.00	Matches:	201
Percent Similarity:	89.45%	Conservative:	28
Best Local Similarity:	78.52%	Mismatches:	27
Query Match:	80.97%	Indels:	0
DB:	14	Gaps:	0

US-09-990-415A-2 (1-261) x US-10-014-338-1 (1-1134)

QY 6 GlyglutProLeuAspIleasnIleGlnGluProArgTyrPaspGlnSerThrPheLeu 25
 |||||
 ::|||
 Db 17 GGAGACTACCAACCAACATTACATCATCAGGAACCTCGATGGGATCAAGCACTTTCATT 76
 |||||

QY 26 GLYARGALAARGHISPHEPETHRVAlThrasPProARgAsnLeuLeuSeRcIyAla 45
 ||||| |
 Db 77 GGACGAGCCATCTTCTTCACTGTAACTGACCCAGGACATCTCTGTAAACCAAGAA 136

QY 46 GlnLeuGIuAlaSerArgasnIleValGlnasnTYrrArgAlaGlyValValThrProGly 65
 |||||:::||||| ||||| ::||| ||||| |||||
 Db 137 CAACTCGAGAGTGCAGAAAATAGTACATGATTACAGCAAGCAATTGTTCTCTCTGGT 196

QY 66 ILeThrGIuSpGIInLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 197 CTTCACGAAATGAATTGTGCACAGCAAGTACATCTATGATTCAGCTTTTCATCTGAC 256

QY 86 ThrGlygluLysvalValleuileglyArgmetSeraIaglnvalPrometasmethr 105
|||||
Db 257 ACTGGTAGAGATGATTTTGATAGGAAGAATGTCAGCCAGGTTCCCATGACATGACC 316

QY 106 ILeThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
|||||
Db 317 ATCACAGGTGTATGATGACGTTTACAGGACTACGCCGGCTGTGCTGTTCGGCAGTGg 376

```
QY      126 ValAsnGlnSerPheasnaIaIleValAsnTyrSerAsnargSergIAspThrProIle 145
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      377 ATTAACCAAGTCCTTCATGCCGTCGTCAATTACACCAACAGAGTGGAGACGCAACCCCTC 436
```

QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
||||| ::

Db 437 ACTGTCATAGAGTTGGGACACAGCTTACGTTTCTGCACACACTGCTGCGGTAGCAACAGCT 496

QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 ||||| ::|||
 Db 497 CTAGACCTCATGCATTGACCAAGCATGTCTCACCACGTATAGAGCTTTTGTTCCTTTT 556

QY 186 AlaAlaValAlaAlaIaAsnCysIleAsnIleProLeuMetArgGlnArgIuLeuGln 205
 |||||
 Db 557 GCTGCCGTAGCTGCTGTAATTGCATTATATTCATTAAATGAGGCAAGGGAACTCAAA 616

Oy	206 ValGlyIleProValAlaaspGluaLagLylnArgLeugLyTyRSeValThraAla	225
Db	617 GTGGCATTCCCGTACAGSATGAGAATGGGAACCGCTTGGGGAGTCGACCAACGCTGC	676

Oy 226 LysGlnGlyIlePheClnValValIleSerArgIleCysMetalIleProAlaMetaIa 245
 ||||| | | | | | | | | | | | | | | | |
 Db 677 AAACAAGCCATCACGCAGTTGTCTGTCACAGATTCTCATGCCAAGCCCCCAATCCCC 736

QY 246 IleProProlLeuMetAspThrLeuGluLysLysAspPheLeuLys 261
||||| :||||| :|||||
Db 737 ATCCCTCCATTCAATATGACACCTTTGGAAAAGCAAGCCGTTTTTTCAC 704

RESULT 5

US-09-990-415A-7
: Sequence 7, Application US/09990415A
: Patent No. US20020165182A1
: GENERAL INFORMATION:

```

; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (125)..(1093)
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: N=A,T,G or C
US-09-990-415A-7

```

Alignment Scores:

Pred. No.:	3.17e-136	length:	2265
Score:	1089.00	Matches:	201
Percent Similarity:	89.45%	Conservative:	28
Best Local Similarity:	78.52%	Mismatches:	27
Query Match:	80.97%	Indels:	0
DB:	10	Gaps:	0

US-09-990-415A-2 (1-261) x US-09-990-415A-7 (1-2269)

Dy 6 GlyGluLeuProLeuAspIleasnIleGlnLuProArgTrpaspGlnSerThrPheLeu 25
 ||||| ::::| ::::| ::::| ::::| :::
Db 131 GGAGAACTACCAACAATTAACATCAAGAACCCTCGATGGGATCAAAGCATTTCATT 190

QY 26 G I Y A R G A I A R G H I S P H E P H E T H V A I T H R A S P R O A R G A S N L E U L E U S E R G I Y A L A 45
 |||||
 Db 191 G G A C G A G C C A T T C T T C A C T G T A C T G A C C C C A G G A C A T T C T G T T A C C A C G A A 250

QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
|||||:::|||||:::|||||:::|||||
Db 251 CAACTCGAGAGTCCGAGAAAATAGTACATGATTCACAGCGAGGAATGTTCCTCCTGGT 310

QY 66 ILEThrglUAspglMLeutprAlgalalystYrrValTYrrAspSeralaphenISProasp 85
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 311 CTTACAGAAATGATTTGCGAGCAAGTACATCTATGATTACAGTTTCACCTTTCCTCCTGAC 370

QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
|||:::|||||
Db 371 ACTGGTAGAGATGATTTTGATAGGAAGAATGTCAGCCAGTTCCCATGAACATGACC 430

QY 106 ILEThrGLYCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTryp 125
|||||
Db 431 ATCACAGGTGTATGATGACGTTTACAGGACTACGCCGGCTGTGCTGTTCCTGGCACTGG 490

```
QY      126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      491 ATTAACCAAGTCCTTCATATGCCGTCGTCAATTACACCACAGAGCTGGAGACGCCACCCCTC 550
```

Qy 146 ThrValArgGlnLeuGlyThrAlaTyrrValSerAlaThrThrGlyAlaValAlaThrAla 165
 ||||| ::||||||||||||||||||||||||||||||||||||||
 Db 551 ACTGTCATAGAGTTGGGACACAGCTTACGTTTCGCAACAACAGTGGTGGCGTAGCAACAGCT 610

```

QY      166 LeuGIleuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
        |||||      ::|||
DB      611 CTAGACCTCATGTCATTGCACCAAGCATGTCTCACCACGTATAGAGCGTTTGTCCCTTT 670

```

QY 186 AlaAlaValAlaAlaIaIaCysIleasnIleProLeuMetArgGlnArgGluLeuGln 205
|||||
Db 671 GCTGCCGTAGCTGCTGAATGTGCATTAATATTCATTAATGAGGCAAGGGAACTCAA 730
|||||

QY 206 ValGlyIleProValAlaaspGluAlaGlyInArgLeuGlyTyrSerValThrAlaAla 225
|||||
Db 731 GTTGACATTCCTCCGTACGGATGAGAAATGGGAACCGCTTGGGGAGTGGCGCAACGCTGGC 790

Db 731 GTTGGCATTCCCGTCACGGATGAGATGGGAAACCGCTTGGGGGAGTCCGGCAACGCTGG 790

```
Db      751 TCGGTGACTGCAGCCAGCAGGGAATCTTCAGGTGGTGAATTCAGAATCTGCATGGCG 810
QY      241 ILeProAlaMetaIaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db      811 ATTCTGCGCATGGCCATCCCACTGATCATGACACTCTGAGAGAAGAAGACTTCCTG 870
QY      261 Lys 261
Db      871 AAG 873

RESULT 3
US-10-037-270-1016
: Sequence 1016, Application US/10037270
: Publication No. US20030104529A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyan
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aidong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yungqing
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhiwei
: APPLICANT: Tillinghast, John
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
: FILE REFERENCE: Polypeptides
: CURRENT APPLICATION NUMBER: US/10/037,270
: CURRENT FILING DATE: 2002-01-04
: PRIOR APPLICATION NUMBER: 09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 1104
: SOFTWARE: pt_fl_genes Version 1.0
: SEQ ID NO 1016
: LENGTH: 1729
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (247)..(1371)
: US-10-037-270-1016

Alignment Scores:
Pred. No.:      5.45e-171      length:      1729
Score:          1345.00      Matches:      261
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             14          Gaps:         0

US-09-990-415A-2 (1-261) x US-10-037-270-1016 (1-1729)
QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db      247 ATGGAAAGCAAATGGGTGAATGCCCTTAGACATCAACATCCAGAACTCGCTGGGAC 306
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db      307 CAAGTACTTTCCTGGGCGAGAGCCCGGCACCTTTTCACTGTACTGATCCTCGAAATCTG 366
QY      41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
```

```
Db      367 CTGCTGTCCGGGGCCACAGCTGGAGACTTCTCGGAACATCGTGCAGAACTACAGGGCCGCC 426
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTyrArgAlaLysTyrValTyrAspSer 80
Db      427 GTGGTAGCCCCAGGGATCACCAGAGACCACTGTGGAGGGCCCAAGTATGTATGACTCC 486
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db      487 GCCTTCCATCCGGACACAGAGGGAGAAAGTGTCTTGATTGGCCGCAATGTACGCCAGGTG 546
QY      101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db      547 CCCATGACATGACCATCAGCTGCGTGCATGCTCACATTCTACAGAGAACCCCAACCGTG 606
QY      121 ValPheThrProlTyrValAlaSerGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db      607 GTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAACACTCCCAACCGCAGT 666
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db      667 GGTGACACTCCCATCACTGTGAGGCGAGCTGGGGACAGCCTATGTGAGTGCACCACTGGA 726
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db      727 GCTGTGGCCAGCGCCCTGGGAGCTGCAACTGCCAAGCACCCTGCCCTTGTCGGC 786
QY      181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db      787 AGATTGTGCCCTTTGACAGCAGTGGCAGCTGCCAAGCACCCTGCCCTTGTCGGC 846
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db      847 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGACAGTGAAGGCTTGCTAC 906
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db      907 TCGGTGACTGCAGCCAAGCAGGGAATCTCCAGGTGTGATTTCAAGATCTGCATGGCG 966
QY      241 ILeProAlaMetaIaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db      967 ATTCTGCGCATGGCCATCCCACTGATCATGACACTCTGAGAGAAGAAGACTTCCTG 1026
QY      261 Lys 261
Db      1027 AAG 1029

RESULT 4
US-10-014-338-1
: Sequence 1, Application US/10014338
: Publication No. US20030092614A1
: GENERAL INFORMATION:
: APPLICANT: Herath, et al.
: TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENAT
: FILE REFERENCE: USES THEREFOR
: CURRENT APPLICATION NUMBER: US/10/014,338
: CURRENT FILING DATE: 2002-05-01
: PRIOR APPLICATION NUMBER: 10/014,338
: PRIOR FILING DATE: 2001-12-10
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1134
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: (1121)..(1122)
: OTHER INFORMATION: where "n" is any nucleotide
: NAME/KEY: misc-feature
: LOCATION: (1125)..(1126)
: OTHER INFORMATION: where "n" is any nucleotide
```


QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
||||| ||| |||||||:::||||| ||| |||
Db 791 AAACAAGCCATCAGCAAGTTGCGTCCAGGATTCATGGCAGCCCTGGCATGGCC 850
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
||||| ||| |||:::||||| ||| |||
Db 851 ATCCCTCCATTCATTATGAACACTTTGGAAAAGAACCTTTTGAAG 898

RESULT 6

US-10-198-846-9767
: Sequence 9767, Application US/10198846
: Publication No. US2003009974A1
: GENERAL INFORMATION:
: APPLICANT: Lille, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steilmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9767
: LENGTH: 2978
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1, 2, 2974, 2975, 2976, 2977, 2978
: OTHER INFORMATION: n - A,T,C or G
US-10-198-846-9767

Alignment Scores:
Pred. No.: 4,88e-136 Length: 2978
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 14 Gaps: 0

US-09-990-415A-2 (1-261) x US-10-198-846-9767 (1-2978)

QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTTrpAspGlnSerThrPheLeu 25
||||| ||| |||:::||||| ||| |||
Db 128 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAAAGCACTTCATT 187
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
||||| ||| |||:::||||| ||| |||
Db 188 GGACGAGCCCAATCATTTCTTCACTGTAAGTGAACCCAGGAACATTTCTGTTAACCAACGAA 247
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
||||| ||| |||:::||||| ||| |||
Db 248 CAACTCGAAGAGTGCAGAAAAATAGTACATGATTACAGGCAAGAAATTGTTCTCTGTTGT 307
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
:::|||||:::|||||:::|||||:::|||||:::|||||
Db 308 CTTACAGAAAATGAATTGTGAGAGCAAAAGTACATCTATGATTCAGCTTTTCATCCTGAC 367
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
||||| ||| |||:::||||| ||| |||
Db 368 ACTGTGAGAAGATGATTTGTATGAGGAAGATGTACAGCCAGGTTCCCATGAACATGACC 427
QY 106 IleThrGlyCysMetLeuThrPheThrArgLysThrProThrValValPheTrpGlnTrp 125
||||| ||| |||:::||||| ||| |||
Db 428 ATCACAGTTGTATGATGACGTTTACAGGACTACGCCGGCTGTCTGTGGCAGTGG 487
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
:::|||||:::|||||:::|||||:::|||||:::|||||

Db 488 ATTAACCACTCCTTCATATGCCGTGTCATTTACACCACAGAGTGGAGACGACCCCTC 547
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
||||| ||| |||:::||||| ||| |||
Db 548 ACTGTCAATGAGTTGGGAACACGTTTACGTTTCTGCACAACTGGTCCGTAGCAACAGCT 607
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
||||| ||| |||:::||||| ||| |||
Db 608 CTAGACTCAATGCATTGACCAAGCATGTCTCACCACCTGATAGACGTTTGTTCCTTT 667
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
||||| ||| |||:::||||| ||| |||
Db 668 GCTGCCGTAGCTGCTGCTAATTGCAATTAATTCATTATGAGGCAAAAGGAACCTCAA 727
QY 206 ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
||||| ||| |||:::||||| ||| |||
Db 728 GTTGGCATTCCTCCGTACAGGATGAGAAATGGGAACCGCTTGGGGAGTCCGCGAACGCTGCG 787
QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
||||| ||| |||:::||||| ||| |||
Db 788 AAACAAGCCATCAGCAAGTTGTCGTCCAGGATTCATTCATGGCAGCCCTGGCATGGCC 847
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
||||| ||| |||:::||||| ||| |||
Db 848 ATCCCTCCATTCATTATGAACACTTTGGAAGAACGCTTTTGAAG 895

RESULT 7

US-09-990-415A-3
: Sequence 3, Application US/09990415A
: Patent No. US20020165182A1
: GENERAL INFORMATION:
: APPLICANT: Pharmacia AB
: TITLE OF INVENTION: Protein Cluster I
: FILE REFERENCE: 00349
: CURRENT APPLICATION NUMBER: US/09/990,415A
: CURRENT FILING DATE: 2001-11-21
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 1061
: TYPE: DNA
: ORGANISM: human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (450)..(680)
US-09-990-415A-3

Alignment Scores:
Pred. No.: 3,37e-125 Length: 1061
Score: 1003.50 Matches: 204
Percent Similarity: 78.16% Conservative: 0
Best Local Similarity: 78.16% Mismatches: 57
Query Match: 74.61% Indels: 1
DB: 10 Gaps: 1

US-09-990-415A-2 (1-261) x US-09-990-415A-3 (1-1061)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTTrpAsp 20
||||| ||| |||:::||||| ||| |||
Db 450 ATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAACCTCGCTGGAC 509
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPheThrValThrAspProArgAsnLeu 40
||||| ||| |||:::||||| ||| |||
Db 510 CAAAGTACTTCTCGGCGACAGAGCCGCACTTTTCACTGTACTGATCCTCGAAATCTG 569
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
||||| ||| |||:::||||| ||| |||
Db 570 CTGCTGTCCGGGGACACAGCTGAAGCTTCTCGGAACATCGTCAGAAC----- 617
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db 617 ----- 617

QY 81 AlaphHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 617 ----- 617
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 618 -----TACAGGAAGACCCCAACCGTG 638
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 639 GTGTTCTGGCAGTGGTGAATCAGTCTTCATGCTTCAATGCTTAACTACTCCAAACCGCAGT 698
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 699 GGTGACACTCCCATCACTGTGAGGCAGCTGGGACAGCCTATGTAGTGCACCACTGGA 758
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 759 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCCTGCGCCCTGTCGGC 818
QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 819 AGATTGTGCTTGGCAGCAGTGGCAGTGCACACTGCATCAACATCCCCCTGATGAGG 878
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 879 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGGAGGTCAAGAGCTTGCGTAC 938
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 939 TCGGTGACTGCAGCAAGCAGGGAATCTTCCAGGTGCTGATTTCAAGAACTTGCATGGCG 998
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 999 ATTCCTGCCATGGCCATCCACCACCTGATGACACTCTGGAGAGAAAGACTTCTCCTG 1058
QY 261 Lys 261
Db 1059 AAG 1061
RESULT 8
US-10-014-338-3
; Sequence 3, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE AN
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (949)..(950)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (979)..(980)
; OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-3
Alignment Scores:
Pred. No.: 6.22e-111 Length: 984
Score: 897.50 Matches: 169
Percent Similarity: 76.26% Conservative: 27

Best Local Similarity: 65.76% Mismatches: 17
Query Match: 66.73% Indels: 44
DB: 14 Gaps: 1
US-09-990-415a-2 (1-261) x US-10-014-338-3 (1-984)
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 17 GGAGAACTACCAACCAACATTAACTAAGAGAACCTCGATGGATCAAGACACTTTCATT 76
QY 26 GlyArgAlaArgHisPheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
Db 77 GGACGAGCCAAATCAATTCTTCACTGTAAGTACCCAGAACATCTCTTAACCAACGAA 136
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 137 CAATCGAGAGTGCAGAGAAATAATAGTACATGATTAACAGCAAGAAATGTTCTCCTGCT 196
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 197 CTACAGAAATGAATTTGTGAGAGCAAGATACATCTATGATTACAGCTTTCATCTCTGAC 256
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 257 ACTGTGAGAGAGATGATTTGTATAGGAAGATGTCAGCCAGGTTCCCATGAACATGACC 316
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 317 ATCAGAGGTGTATGATGAGCTTTTACAGGACTACGCCGCTGTGCTGTGCGCAGTGG 376
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 377 ATTAACCAAGTCTTCAATGCCGCTGCTCAATTAACCAACAGAAAGTGAGAGCACCCCTC 436
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 437 ACTGTCATGAGTTGGGAACAGCTTACGTTCTGTAACAACTGTCGCCGTAGCAACAGCT 496
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 497 CTAGGACTCATGATGATGACCAAGCATGTCTCACCACTGATAGAGCTTTGTTCCTTT 556
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db 557 GCTGCCGTAGCTGCTGCTAATGCTAATTAATTCATTATGAGGCAA----- 604
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 604 ----- 604
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 605 -----AGC 607
QY 246 -IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 608 CATCCCTCATTCATTATGAACACTTGGAAAAAGAAAGCCCTTTTGAAG 656
RESULT 9
US-09-990-415a-5
; Sequence 5, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1567
; TYPE: DNA

ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1015)
US-09-990-415A-5

Alignment Scores:
Pred. No.: 3.14e-91 Length: 1567
Score: 754.00 Matches: 150
Percent Similarity: 72.20% Conservative: 37
Best Local Similarity: 57.92% Mismatches: 72
Query Match: 56.06% Indels: 0
DB: 10 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-990-415A-5 (1-1567)

QY 3 SerlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTrpAspGlnSer 22
DB 41 AGCAAGATGAGGCTGACCTGTCTGCTTACATCGATGCCCGCGTGGGACCGCCG 100
QY 23 ThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 42
DB 101 ACCTTCCTGGGAGAGTGAAGCACTTCCTAACATCAGGACCCCGGCACTGTCTTGTGA 160
QY 43 SerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValVal 62
DB 161 TCTGAGCGGGAGCTGGAGCTGGGCCAAGGTGATGCTGAGAGAAGACAGATGGGGTTGTG 220
QY 63 ThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPhe 82
DB 221 CCCCAGGACCCCAAGTGAGAGCAGCTGTGTATGCCAAAGAGCTGATGCTGGCCTTC 280
QY 83 HisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValPromet 102
DB 281 CACCCGACACTGGGGAGAGAAGATGAATGTCAATCGGCGCATGTCTTCCAGCTTCCCTGGC 340
QY 103 AsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPhe 122
DB 341 GGCATGATCATCACGGGCTTCATGCTCCAGTCTACAGAGCATGGCGGGTGATCTTC 400
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAsp 142
DB 401 TGGCAGTGGGTGAACACAGTCTCTCAATGCCCTTAGTCAACTACACCAACAGAAATGCGGCT 460
QY 143 ThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaVal 162
DB 461 TCCCCACATCATCAGGACAGATGGCCCTTCTTCTACTCTCACAGCCACCACTGCTGTG 520
QY 163 AlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPhe 182
DB 521 GCCACGGCTGTGGCATGAACATGTTGACAAAGAACGCCGCCCTGTGGTGGCGCTGG 580
QY 183 ValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArg 202
DB 581 GTGCCCTTGGCGCTGTGGCTGGCGGCTACTGTCTCAATATCCCATGATGCGACAGAGG 640
QY 203 GluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerVal 222
DB 641 GAGCTCATAAAGGAATCTGCGTGAAGACAGGAATGAAATGAGATGTGTCATTCGCCG 700
QY 223 ThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetalIlePro 242
DB 701 AGAGCTGGGCGCATAGGCATCACCCAAGTGTATTCTCGGATCACCATGTCACTCCT 760
QY 243 AlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
DB 761 GGGATGATCTTGTCTGCCAGTCAATGAAAGGCTTGAGAAATTGCACCTTCATGCAG 817

RESULT 10
US-10-098-841-64
; Sequence 64, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
FILE REFERENCE: 784CIP2
CURRENT APPLICATION NUMBER: US/10/098,841
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pt_fl_genes version 1.0
SEQ ID NO 64
LENGTH: 2376
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (359)..(1327)
US-10-098-841-64

Alignment Scores:
Pred. No.: 2.1e-90 Length: 2376
Score: 750.00 Matches: 149
Percent Similarity: 72.20% Conservative: 38
Best Local Similarity: 57.53% Mismatches: 72
Query Match: 55.76% Indels: 0
DB: 13 Gaps: 0

US-09-990-415A-2 (1-261) x US-10-098-841-64 (1-2376)

QY 3 SerlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTrpAspGlnSer 22
DB 353 AGCAAGATGAGGCTGACCTGTCTGCTTACATCGATGCCCGCGTGGGACCGCCG 412
QY 23 ThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 42
DB 413 ACCTTCCTGGGAGAGTGAAGCACTTCCTAACATCAGGACCCCGGCACTGTCTTGTGA 472
QY 43 SerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValVal 62
DB 473 TCTGAGCGGAGCTGAGCTGGGCCAAGGTGATGTGAGAGAAGACAGATGGGGTTGTG 532
QY 63 ThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPhe 82
DB 533 CCCCAGGACCCCAAGTGAGACAGCTGTGTATGCCAAGAAGCTGTATGACTCGGCTTC 592
QY 83 HisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValPromet 102
DB 593 CACCCGACACTGGGAGAGAAGATGAATGTATCGGGCGCATGTCTTCCAGCTTCCCTGGC 652
QY 103 AsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPhe 122
DB 653 GGCATGATCATCACGGGCTTCATGCTCAGTTCTACAGAGCATGCCGCGGTGATCTTC 712
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAsp 142

```
Db      713 TGGCAGTGGGTGAACAGTCCCTTCATGCGCTTACTCACTACACCAAGGAATGCGGCT 772
QY      143 ThrProIleThrValArgInLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaVal 162
Db      773 TCCCCACATCAGTCACAGCAGATGGCCCTTCTCTACTTCACAGCCACCACTGCTGTG 832
QY      163 AlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPhe 182
Db      833 GCCACGGCTGTGGGCATGAACATGTTGACAAAGAACGCCGCCCTTGTGGCCGCTGG 892
QY      183 ValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArg 202
Db      893 GTGCCCTTGGCGCTGTGGCTGCGGCTAACTGTGCAATATCCCCATGATGCGACAGCAG 952
QY      203 GluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerVal 222
Db      953 GAGCTCATAAAGGAATCTGCGTGAAGGACAGGAATGAATAATGAGATTGTCATCCCGG 1012
QY      223 ThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIlePro 242
Db      1013 AGAGCTGGCGGCATAGGCATCACCCAGTAGTATTCTCGGATCACCATGTCACTCCT 1072
QY      243 AlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db      1073 GGGATGATCTTGTCTGCCAGTCATCATGGAAGGCTTGAGAAATTGCACTTCATGCAG 1129
```

RESULT 11

```
US-09-822-849A-490
; Sequence 490, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukote, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 490
; LENGTH: 2458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-490
```

Alignment Scores:

```
Pred. No.: 2.22e-90 Length: 2458
Score: 750.00 Matches: 149
Percent Similarity: 72.20% Conservative: 38
Best Local Similarity: 57.53% Mismatches: 72
Query Match: 55.76% Indels: 0
DB: 9 Gaps: 0
```

US-09-990-415A-2 (1-261) x US-09-822-849A-490 (1-2458)

```
QY      3 SerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTTPAspGlnSer 22
Db      75 AGCAAGATGGAGGCTGACCTGTCTGGCTTTAACATCGATGCCCGCTTGGAGCCAGCGC 134
QY      23 ThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 42
Db      135 ACCTTCTTGGGGAGAGTGAAGCACTTCTTAACATCACGAGCCCGCAGCTGTCTTTGTA 194
```

```
QY      43 SerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValVal 62
Db      195 TCTGAGCGGGAGCTGGCACTGGGCCAAGGTGATGTTGGAAGAAGCAGCATGGGGTTGTG 254
QY      63 ThrProGlyIleThrGluAspGlnLeuTTPArgAlaLysTyrValTyrAspSerAlaPhe 82
Db      255 CCCCAGGCACCCCAAGTGGAGCAGCAGCTGTGTATGCCAAAAGCTGTATGACTCGGCTTC 314
QY      83 HisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPromet 102
Db      315 CACCCGACACTGGGGAGAGAAGATGAATGTCAATCGGGCGCATGTCTTCCAGCTTCCTGGC 374
QY      103 AsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValPhe 122
Db      375 GGCATGATCATCAGCGGCTTTCATGCTCAGTTCTACAGACGATGCCGGGCTGATCTTC 434
QY      123 TRPGLNTPValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAsp 142
Db      435 TGGCAGTGGGTGAACCAAGTCCCTTCAATGCCCTTAGTCAACTACACCAACAGGAATGCGGCT 494
QY      143 ThrProIleThrValArgInLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaVal 162
Db      495 TCCCCACATCAGTCACGAGATGGCCCTTCTCTACTTCACAGCCACCACTGCTGTG 554
QY      163 AlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPhe 182
Db      555 GCCACGGCTGTGGGCATGAACATGTTGACAAAGAACGCCGCCCTTGTGGCGCGCTGG 614
QY      183 ValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArg 202
Db      615 GTGCCCTTGGCGCTGTGGCTGCGGCTAACTGTGCAATATCCCCATGATGCGACAGCAG 674
QY      203 GluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerVal 222
Db      675 GAGCTCATAAAGGAATCTGCGTGAAGGACAGGAATGAATAATGAGATTGTCATTCCCGG 734
QY      223 ThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIlePro 242
Db      735 AGAGCTGGGGCCATAGGCATCACCCAGTAGTATTCTCTCGGATCACCATGTCACTCCT 794
QY      243 AlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db      795 GGGATGATCTTGTCTGCCAGTCATCATGGAAGGCTTGAGAAATTGCACTTCATGCAG 851
```

RESULT 12

```
US-10-206-901B-41
; Sequence 41, Application US/10206901B
; Publication No. US20030100540A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, ZHONGHUA
; APPLICANT: DUBOIS, RAYMOND
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES
; FILE REFERENCE: VBLT:012US
; CURRENT APPLICATION NUMBER: US/10/206,901B
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/308,370
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-901B-41
```

Alignment Scores:

```
Pred. No.: 1.34e-85 Length: 670
Score: 708.00 Matches: 135
Percent Similarity: 88.62% Conservative: 13
Best Local Similarity: 80.84% Mismatches: 19
Query Match: 52.64% Indels: 0
DB: 14 Gaps: 0
```


US-09-990-415A-2 (1-261) x US-10-206-901B-41 (1-670)

QY 95 ArgMetSerAlaGlnValPrometAsnMetThrIleThrGlyCysMetLeuThrPheTyr 114
Db 3 AGAATGTCAGCCCGAGTTCCTCCATGAACATGACCATCACAGGTTGTATGATGACGTTTAC 62
QY 115 ArgIleThrProThrValValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleVal 134
Db 63 AGGACTACGCCGGCTGTGCTGTCTGGCAGTGATTAACCAAGCTTCAATGCCGTGTC 122
QY 135 AsnTyrSerAsnArgSerGlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyr 154
Db 123 AATTAACCAACAGAGAGTGAGACGCCACCCCTCACTGTCAATGAGTTGGGAACAGCTTAC 182
QY 155 ValSerAlaThrThrGlyAlaValAlaThrAlaLeuGlyLeuIleSerLeuThrIleHis 174
Db 183 GTTCTGCAACAACACTGTGTCCTGACCAACAGCTCTAGGACTCAATGCAATTGACCAAGCAT 242
QY 175 LeuProProLeuValGlyArgPheValProPheAlaAlaAlaAlaAlaAsnCysIle 194
Db 243 GTCACACCACTGATAGACGTTTGTCTTCCCTTGTGCTGCCGTAAGCTGCTGAATTGCATT 302
QY 195 AsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIleProValAlaAspGluAla 214
Db 303 AATATTCATTAAATGAGGCAAGGAACTCAAGTTGGCATCCCGTCACGAGTAGAAT 362
QY 215 GlyGlnArgLeuGlyTyrSerValThrAlaAlaAlaGlnGlyIlePheGlnValIle 234
Db 363 GGGAAACCGCTTGGGAGTGGCGGAACGCTGCGAAACAAGCCATCACGCAAGTGTGCTG 422
QY 235 SerArgIleCysMetAlaIleProAlaMetAlaIleProProLeuIleMetAspThrLeu 254
Db 423 TCCAGGATTCCTCATGGCAGCCCTGGCATGGCCATCCCTCCATTCAATTATGAACACTTG 482
QY 255 GluIleAspPheLeuIle 261
Db 483 GAAAGAAAGCCTTTTGAAG 503

RESULT 13

US-09-918-995-23576
; Sequence 23576, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918, 995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23576
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23576

Alignment Scores:

Pred. No.: 1.79e-69 Length: 489
Score: 587.00 Matches: 101
Percent Similarity: 92.31% Conservative: 19
Best Local Similarity: 77.69% Mismatches: 10
Query Match: 43.64% Indels: 0
DB: 11 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-918-995-23576 (1-489)

QY 7 GluLeuProLeuAspIleAsnIleGlnIleProArgTrpAspGlnSerThrPheLeuGly 26
Db 98 GAACTACCAACCAACATTAACATCAAGGAACCTGATGGATCAAAACACTTTCATTGGA 157

QY 27 ArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAlaGln 46
Db 158 CGAGCCAAATCATTTCTTCACTGTAACGACCCAGAGACATCTCTTAACCAACAGAACAA 217
QY 47 LeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGlyIle 66
Db 218 CTCGAGAGTCGAGAAAAATAGTACATGATTACAGGCAAGGAATGTTCTCCTGCTT 277
QY 67 ThrGlnAspGlnLeuTrpArgAlaIleTyrValTyrAspSerAlaPheHisProAspThr 86
Db 278 ACAGAAATGAATTTGTGAGAGCAAAAGTACATCTATGATTCAGCTTTTCATCCTGACACT 337
QY 87 GlyGlnValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThrIle 106
Db 338 GTGAGAGATGATTTTGTATGAGAGATGTCAGCCAGGTTCCATGAACATGACCATC 397
QY 107 ThrGlyCysMetLeuThrPheTyrArgIleThrProThrValValPheTrpGlnTrpVal 126
Db 398 ACAGGTTCTATGATGACGTTTTCAGAGACTACGCCGGCTGTGCTGTCTGCGCAGTGATT 457
QY 127 AsnGlnSerPheAsnAlaIleValAsnTyr 136
Db 458 AACCAAGTCTTCAATGCCGTGTCATTTAC 487

RESULT 14

US-09-814-353-20153
; Sequence 20153, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814, 353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20153
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3341, 3342
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20153

Alignment Scores:

Pred. No.: 2.61e-49 Length: 3342
Score: 447.50 Matches: 90
Percent Similarity: 57.60% Conservative: 54
Best Local Similarity: 36.00% Mismatches: 98
Query Match: 33.27% Indels: 8
DB: 12 Gaps: 2

US-09-990-415A-2 (1-261) x US-09-814-353-20153 (1-3342)

QY 16 GluProArgTrpAspGlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThr 35

[illegible]

OTHER INFORMATION: n - A,T,C or G
US-09-918-995-5675.

Alignment Scores:	
Pred. No.:	1.7e-38
Score:	358.00
Percent Similarity:	84.62%
Best Local Similarity:	79.12%
Query Match:	26.62%
DB:	11
Length:	522
Matches:	72
Conservative:	5
Mismatches:	14
Indels:	0
Gaps:	0

US-09-990-415A-2 (1-261) x US-09-918-995-5675 (1-522)

QY	171	LeuThrLysHisIleuProPheValGlyArgPheValProPheAlaIaValAlaIa	190
Dd	3	TTGACCAAGCATGTCTCACCACCTGATAGGACGTTTGTGCCCTTGCGCCGTAGCTGCT	62
QY	191	AlaAsnCysIleasnIleProIeuMetArgGlnArgGluLeuGlnValGlyIleProVal	210
Dd	63	GCTAATTGCATTAAATATTCATTAAAGAGCAAAGGAAGACTCAAAAGTTGGCATGCCGTC	122
QY	211	AlaaspGluAlaGlyGlnArgLeuGlyTyrSerValThraIaAlaLysGlnGlyIlePhe	230
Dd	123	ACGGATGAGAATGGGAACCGCTTGGGGGAGTCGGCGAACGCTCGGAACAAGCCATCACG	182
QY	231	GlnValValIleSerArgIleCysMetalaIaIleProAlaMetalaIaIleProPheIle	250
Dd	183	CAAGTGTGCTGCTCCAGGATTTCTCATGCGACGCCCTGGCATGGCCATCCCTCCATTTCATT	242
QY	251	MetAspThrLeuGluLysLysAspPheLeuLys	261
Dd	243	ATGAACACTTTGGAAAAGAAAGCCCTTTTGAAG	275

Search completed: October 16, 2003, 17:40:34
Job time : 269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 15:58:12 ; Search time 2075 Seconds
(without alignments)
3057.093 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPLDINIQEPRWD.....PAMAIPLIMDTLEKKDFLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool.p/US09990415/runat_16102003_073807_9449/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdl -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415.ecgn_1_1_2874_erunat_16102003_073807_9449 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estfun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vr1:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1341	99.7	1201	13	BX376574	BX376574 BX376574
2	1333	99.1	869	13	BQ223306	BQ223306 AGENCOURT
3	1325	98.5	975	13	BQ933752	BQ933752 AGENCOURT
4	1312	97.5	998	12	BM473966	BM473966 AGENCOURT
5	1305	97.0	1201	13	BX334089	BX334089 BX334089
6	1291	96.0	1076	12	BM925231	BM925231 AGENCOURT
7	1290	95.9	995	13	BX355800	BX355800 BX355800
8	1283	95.4	898	14	CB203758	CB203758 AGENCOURT
9	1283	95.4	2682	11	AK034514	AK034514 Mus muscu
10	1216	90.4	963	13	BQ959114	BQ959114 AGENCOURT
11	1212	90.1	934	13	BQ506082	BQ506082 AGENCOURT
12	1196	88.9	1040	12	BM548674	BM548674 AGENCOURT
13	1189	88.4	961	12	BG775086	BG775086 602650054
14	1180	87.7	862	13	BQ425697	BQ425697 AGENCOURT
15	1166	86.7	731	14	CA317398	CA317398 UI-M-FW0-
16	1166	86.7	811	12	BI818854	BI818854 603037323
17	1166	86.7	870	12	BM047107	BM047107 603627441
18	1162	86.4	764	14	CA317392	CA317392 UI-M-FW0-
19	1153	85.7	715	14	CA316261	CA316261 UI-M-FW0-
20	1148.5	85.4	868	10	BE780804	BE780804 601469693
21	1121	83.3	1008	10	BF569711	BF569711 602186237
22	1114	82.8	755	12	BI517760	BI517760 603042003
23	1105	82.2	727	14	CD354095	CD354095 UI-M-GM0-
24	1104.5	82.1	1071	10	BF569784	BF569784 602185719
25	1090.5	81.1	758	10	BF346247	BF346247 602018434
26	1089	81.0	1201	13	BX363487	BX363487 BX363487
27	1084.5	80.6	746	12	BI752914	BI752914 603028557
28	1077.5	80.1	2326	11	AK089985	AK089985 Mus muscu
29	1076	80.0	1080	9	AL583707	AL583707 AL583707
30	1073	79.8	869	14	CA489784	CA489784 AGENCOURT
31	1072	79.7	830	13	BU408998	BU408998 603159016
32	1071.5	79.7	666	14	CA319631	CA319631 UI-M-FW0-
33	1069	79.5	864	13	BQ435963	BQ435963 AGENCOURT
34	1059	78.7	1228	11	AK012650	AK012650 Mus muscu
35	1053	78.3	1016	14	BY711785	BY711785 BY711785
36	1048.5	78.0	956	12	BM928783	BM928783 AGENCOURT
37	1048	77.9	955	13	BU254199	BU254199 603341708
38	1041	77.4	791	13	BU355128	BU355128 603474791
39	1039	77.2	786	12	BI152915	BI152915 602918086
40	1036	77.0	1004	12	BM471040	BM471040 AGENCOURT
41	1033	76.8	676	12	BM944910	BM944910 UI-M-EH0P
42	1032.5	76.8	782	14	CB234292	CB234292 AGENCOURT
43	1032	76.7	655	10	BB626574	BB626574 BB626574
44	1032	76.7	793	13	BU355445	BU355445 603473920
45	1017	75.6	944	12	BI903936	BI903936 603166338

ALIGNMENTS

RESULT 1
BX376574
LOCUS BX376574 1201 bp mRNA EST 08-MAY-2003
DEFINITION BX376574 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
CDNA clone CS0DD001YF06 5-PRIME, mRNA sequence.
ACCESSION BX376574
VERSION BX376574.1 GI:30448513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)

AUTHORS L1,W.B., Gruber,C., Jesse,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL unpublished
 COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DD001DC03QPI&cluster=5549.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DD001DC03QPI.
 Location/Qualifiers

FEATURES

source

1.1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD001YF06"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 268 a 306 c 347 g 250 t 30 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.41e-147 Length: 1201
 Score: 1341.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0
 Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 99.70% Indels: 0
 DB: 13 Gaps: 0

US-09-990-415A-2 (1-261) x BX376574 (1-1201)

OY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
 Db 113 ATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGAACCTCGCTGGGAC 172
 OY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 Db 173 CAAAGTACTTTCCTGGGGCAGAGCCCGCAGCTTTTCTCACTGTTACTGATCCTCGAAATCTG 232
 OY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyArgAlaGly 60
 Db 233 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGACAGACTACAGGGCCGGC 292
 OY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyValTyrAspSer 80
 Db 293 GTGGTGACCCCGAGGATCACCGAGACCACTGTGAGGGCCCAAGTATGTGTATGACTCC 352
 OY 81 AlaPheHisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnVal 100
 Db 353 GCCCTTCATCCGGAGACACAGGGAGAGGTGTCTGATGGCCCGCATGTAAAGCCAGGTG 412
 OY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyArgLysThrProThrVal 120
 Db 413 CCCATGAACATGACCATCATCTGGCTGCTCATCATTTCTACAGAGAACCCCAACCGTG 472
 OY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTySerAsnArgSer 140
 Db 473 GTGTCTGTGGCAGTGGGTGAATCAGTCTTCAATGCCATTTGTTAACTACTCCAAACCGCAGT 532
 OY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyValSerAlaThrThrGly 160
 Db 533 GGTGACACTCCCATCATCTGTGAGGCGAGCTGGGGACAGCCATATGTGAGTGCCACCACTGGA 592
 OY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180

Db 593 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCAACCAAGCACCTGCCCTTGCTGGC 652
 OY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 Db 653 AGATTTGTGCCCTTTCAGACAGTGCGAGCTGCCAAGTGCATCAACATCCCTGTATGAGG 712
 OY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyLysArgLeuGlyTyr 220
 Db 713 CAGAGAGAGCTGACAGTGGGCATCCCGGTGGCTGATGAGGACAGTACAGAGCTTGCTAC 772
 OY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 Db 773 TCGGTGACTGCAGCCAGCAGGAGATCTTCAGGTGGTGAATTCAGAAATCTGCATGGCG 832
 OY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 Db 833 ATTCCTGCCATGGCCATCCCAACCACTGATGACACTCTGGAGAGAAGACTTCTCTG 892
 OY 261 Lys 261
 Db 893 AAG 895

RESULT 2 869 bp mRNA linear EST 02-MAY-2002
 BQ223306 AGENCOURT_7518440 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:606572
 LOCUS 5', mRNA sequence.
 DEFINITION BQ223306 BQ223306.1 GI:20404706
 ACCESSION BQ223306
 VERSION BQ223306.1 GI:20404706
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 869)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM13344 row: k column: 21
 High quality sequence stop: 688.

FEATURES

source

1.869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:606572"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

BASE COUNT 193 a 244 c 255 g 175 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 7.79e-147 Length: 869
 Score: 1333.00 Matches: 259
 Percent Similarity: 99.62% Conservative: 1
 Best Local Similarity: 99.23% Mismatches: 1
 Query Match: 99.11% Indels: 0

DB:	13	Gaps:	0
US-09-990-415A-2 (1-261) x BQ223306 (1-869)			
QY	1	MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp	20
Db	37	ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGAC	96
QY	21	GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu	40
Db	97	CAAGTACTTCTCTGGGCGAGAGCCCGGCACTTTTCACTGTTACTGATCCTCGAAATCTG	156
QY	41	LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrrArgAlaGly	60
Db	157	CTGCTGTCCGGGCACAGCTGGAAGCTTCTCGAACATCGTGACAGAACTACAGGGCCGCG	216
QY	61	ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrrValTyrrAspSer	80
Db	217	GTGTGACCCAGGAGATCACCGAGACCAAGTGTGAGGGCCAAGTATGTATGACTCC	276
QY	81	AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal	100
Db	277	GCCTTCCATCCGGACACACAGGGAGAGGTGCTCTGATTGGCCCATGTACGCCACAGTG	336
QY	101	PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrrArgLysThrProThrVal	120
Db	337	CCCATGAACATGACCATCACTGCTGCATGCTCACATTCTACAGGAAGCCCAACCGTG	396
QY	121	ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrrSerAsnArgSer	140
Db	397	GTGTTCTGGCAGTGGGTGAATCAGTCTCTCAATGTCATTGTTAACTACTCCAACCGCAGT	456
QY	141	GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrrValSerAlaThrThrGly	160
Db	457	GGTGACACTCCCATCACTGTGAGCAGCTGGGGACAGCCTATGTAGTGGCACCACTGGA	516
QY	161	AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly	180
Db	517	GCTGTGGCCAGGCCCCCTGGACTCAAAATCCCTCACCAAGCACCTGCCCTTGGTCGGC	576
QY	181	ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg	200
Db	577	AGATTGTGCCCCCTTTGCAGCAGTGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAG	636
QY	201	GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr	220
Db	637	CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGGAGGTCAGAAAGCTTGGCTAC	696
QY	221	SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla	240
Db	697	TCGGTGAAGTGCAGCAAGCAGGGAATCTTCAGGTGATTTCAAGAACTCTGCATGGCG	756
QY	241	IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu	260
Db	757	ATTCCTGCCATGGCCATCCCAACAGCTGATGAGACACTCTGAGAGAAGAAGACTTCCTG	816
QY	261	Lys 261	
Db	817	AAG 819	
RESULT 3			
LOCUS	BQ933752	975 bp	mRNA linear EST 21-AUG-2002
DEFINITION	AGENCOURT_8837545 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6424845		
ACCESSION	BQ933752		
VERSION	BQ933752.1 GI:22349135		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 975)		

AUTHORS		NIH-MGC http://mgc.ncl.nih.gov/ .	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: DCTD/DTP/Gazdar CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2605 row: k column: 22 High quality sequence stop: 646. Location/Qualifiers	
FEATURES		1. .975	
source		/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6424845" /tissue_type="large cell carcinoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_18" /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	
BASE COUNT		225 a	277 c 267 g 199 t 7 others
ORIGIN			
Alignment Scores:			
Pred. No.:		8.11e-146	Length: 975
Score:		1325.00	Matches: 256
Percent Similarity:		99.61%	Conservative: 2
Best Local Similarity:		98.84%	Mismatches: 1
Query Match:		98.51%	Indels: 0
DB:		13	Gaps: 0
US-09-990-415A-2 (1-261) x BQ933752 (1-975)			
QY	1	MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp	20
Db	30	ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGAC	89
QY	21	GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu	40
Db	90	CAAGTACTTCTCTGGGCGAGAGCCCGGCACTTTTCACTGTTACTGATCCTCGAAATCTG	149
QY	41	LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrrArgAlaGly	60
Db	150	CTGCTGTCCGGGCACAGCTGGAAGCTTCTCGGAACATCGTGACAGAACTACAGGGCCG	209
QY	61	ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrrValTyrrAspSer	80
Db	210	GTGTGACCCAGGAGATCACCGAGACCAAGTGTGAGGGCCAAGTATGTATGACTCC	269
QY	81	AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal	100
Db	270	GCCTTCCATCCGGACACAGGGGAGAAAGTGTCTGATTGGCCGCATGTACGCCACAGTG	329
QY	101	PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrrArgLysThrProThrVal	120
Db	330	CCCATGAACATGACCATCACTGCTGCATGCTCACATTCTACAGGAAGACCCCAACCGTG	389
QY	121	ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrrSerAsnArgSer	140
Db	390	GTGTTCTGGCAGTGGGTGAATCAGTCTCTCAATGTCATTGTTAACTACTCCAACCGCAGT	449
QY	141	GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrrValSerAlaThrThrGly	160

```

Db      450 GGTGACACTCCCATCACTGTGAGGCAGCTGGGAGACAGCCTATGTGAGTGCCACCACCTGGA 509
QY      161 AlavalathralaleuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db      510 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAGCACCCTGCCCCCTTGTCGGC 569
QY      181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db      570 AGATTGTGCTCTTTGACAGCAGTGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAGG 629
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db      630 CAGAGAGAGCTGCAGGTGGGATCCCGGTGGCTGATGAGGCAGGTGAGAAGCTTGCTAC 689
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
Db      690 TCGGTGACTGCAGCCAAAGCAGGGAATCTCCAGGTGTGATTCAGAATCTGCATGGCG 749
QY      241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPhe 259
Db      750 ATTCCCTGCATGGCCATCCACCACTGATGACACTCTGAGAGAAGAAGACTTC 806

RESULT 4
BM473966      998 bp      mRNA      linear      EST 05-FEB-2002
LOCUS      AGENCOURT.6489827 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5521420
ACCESSION      5', mRNA sequence.
VERSION      BM473966
KEYWORDS      BM473966.1 GI:18523008
SOURCE      EST.
ORGANISM      Homo sapiens (human)
HOMO SAPIENS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 998)
AUTHORS      NIH-MGC http://mgs.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Life Technologies, Inc.
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LHAM12187 row: a column: 05
      High quality sequence stop: 652.

FEATURES
    source
        1..998
            location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5521420"
                /tissue_type="leiomyosarcoma"
                /lab_host="DH108 (phage-resistant)"
                /clone_lib="NIH_MGC_71"
                /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 2.1 kb."
BASE COUNT      228 a      292 c      279 g      198 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      2.89e-144      Length:      998
Score:      1312.00      Matches:      259
Percent Similarity:      98.86%      Conservative:      1
Best Local Similarity:      98.48%      Mismatches:      1
Query Match:      97.55%      Indels:      2
DB:      12      Gaps:      0

```

```

US-09-990-415A-2 (1-261) x BM473966 (1-998)
QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIleProArgTrpAsp 20
Db      43 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGCTGGGAC 102
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIleThrAspProArgAsnLeu 40
Db      103 CAAAGTACTTTCTCTGGGCAGAGCCCGGCACTTTTCACTGTTACTGATCTCGAAATCTG 162
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db      163 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGACAGACTACAGGGCCGGC 222
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db      223 GTGTGACCCACAGGATCACCAGACCACTGTGAGGGCCAAAGTATGTATGACTCC 282
QY      81 AlaPheHisProAspThrGlyGlyLysValIleuIleGlyArgMetSerAlaGlnVal 100
Db      283 GCCTTCCATCCGACACAGAGGAGAGAGTGTCTGATTTGGCCGCAATGTACGCCAGGTG 342
QY      101 PrometasMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db      343 CCCATGAACATGACCATCACTGGCTGCATGCTACATTTCTACAGAGAAGACCCCAACCGTG 402
QY      121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db      403 GTGTCTGCGCAGTGGGTGAATCAGTCTTCAATGCCATTGTTAATCTCAACCCGACGT 462
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db      463 GGTGACACTCCCATCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 522
QY      161 AlavalathralaleuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db      523 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGTCGGC 582
QY      181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db      583 AGATTGTGCCCTTTGACAGCAGTGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAGG 642
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db      643 CAGAGAGAGCTGCAGGTGGGCAATCCGGTGGCTGATGAGGCAGGTGAGAAGCTTGCTAC 702
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db      703 TCGGTGACTGCAGCCAAAGCAGGGAATCTTCCAGGTGGGTGATTTCAAGAATCTGCATGG 762
QY      240 IalIleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPhe 260
Db      763 CGATTCTGCTCCATGNCATCCACCACTGATCATGACACTCTGAGAGAAGAAGACTTCC 822
QY      260 eulys 261
Db      823 TGAAG 827

RESULT 5
BX334089      1201 bp      mRNA      linear      EST 02-MAY-2003
LOCUS      BX334089 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION      cDNA clone CS0DD006YP08 5'-PRIME, mRNA sequence.
ACCESSION      BX334089
VERSION      BX334089.1 GI:30341296
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1201)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization

```

JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DD006DH04QPI&cluster=5549.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DD006DH04QPI.

FEATURES
 source
 1.1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YP08"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the PCMVSPORT 6 vector. Library was normalized."

BASE COUNT 266 a 316 c 345 g 245 t 29 others

ORIGIN

Alignment Scores:
 Pred. No.: 2.54e-143 Length: 1201
 Score: 1305.00 Matches: 258
 Percent Similarity: 98.85% Conservative: 0
 Best Local Similarity: 98.85% Mismatches: 3
 Query Match: 97.03% Indels: 1
 DB: 13 Gaps: 0

US-09-990-415A-2 (1-261) x BX334089 (1-1201)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 |||||||
 Db 88 ATGGAAGCAAAATGGGTGARTGGCTTAGACATCAACATCCAGAACCTCGCTGGGAC 147
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 |||||||
 Db 148 CAAAGTACTTCTCTGGGCGAGAGC-CGGCACTTTTCACTGTACTGATCTCGARATCTG 206
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 |||||||
 Db 207 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGAACATCTGTGCAGAACTACAGGCGCGC 266
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPArgAlaAlaLysTyrValTyrAspSer 80
 |||||||
 Db 267 GTGGTGACCCCGAGGATCACCGAGGACAGCTGTGGAGGGGCCAAGTATGTGTAGACTCC 326
 QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 |||||||
 Db 327 GCCTTCCATCCGAGACAGGGGAGAAGGTGCTGCTGATTTGGCCGCATGTACAGCCAGGTG 386
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 |||||||
 Db 387 CCCATGAACATGACATCACTGGCTGCATGCTCACATTTCTACAGGAAGACCCCAACCGTG 446
 QY 121 ValPheTyrGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 |||||||
 Db 447 GTGTCTTGGCAGTGGGGAATCAGTCTTCAATGCCATGTGTTAACTACTCCAACCGCAGT 506
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 |||||||
 Db 507 GGTGACACTCCCATCTACTGTGAGGACAGGAGGAGGACAGCCTATGTGAGTGCCACACTGGA 566
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 |||||||
 Db 567 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACACTGCCCCCTGTGGCGC 626

QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 |||||||
 Db 627 AGATTGTGCCCTTTGCAGACAGATGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAGG 686
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 |||||||
 Db 687 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGAGGCTCAGAGGCTTGCTTAC 746
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 |||||||
 Db 747 TCGGTGACTGCAGCCAAAGCAGGGAATCTTCCAGGTGGTGTGATTTCAAGAATCTGCATGGCG 806
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 |||||||
 Db 807 ATTCCCTGCCATGGCCATCCCACTGATCATGACACATCTGGAGAGAAGAAGACTTCTCTG 866
 QY 261 Lys 261
 |||
 Db 867 AAG 869

RESULT 6
 BM925231
 LOCUS BM925231
 DEFINITION AGENCOURT_6627462 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5762556
 5', mRNA sequence.
 ACCESSION BM925231 GI:19375610
 VERSION BM925231.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1076)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12812 row: p column: 13
 High quality sequence start: 4
 High quality sequence stop: 664.

FEATURES
 source
 1.1076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5762556"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC library."

BASE COUNT 230 a 316 c 314 g 216 t

ORIGIN

Alignment Scores:
 Pred. No.: 9.77e-142 Length: 1076
 Score: 1291.00 Matches: 255

Percent Similarity: 99.22% Conservative: 1
 Best Local Similarity: 98.84% Mismatches: 2
 Query Match: 95.99% Indels: 1
 DB: 12 Gaps: 0

US-09-990-415A-2 (1-261) x BM925231 (1-1076)

```

QY      4 LysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThr 23
Db      10 CAATGGGTGAATGCTTTAGACATCAC-ATCCAGGAACCTCGTGGGACCAAGTACT 68
QY      24 PheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSer 43
Db      69 TTCCTGGGACAGAGCCGGGCACTTTTTCACGTGTACTGATCCTCGAAATCTGCTGTGCC 128
QY      44 GlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThr 63
Db      129 GGGGACAGCTGGAGCTTCTCGGAACATCGTCAGAACTACAGGCGCGGTGTGACC 188
QY      64 ProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHis 83
Db      189 CCAGGATCACCAGGAGACCACTGTGAGAGGGCCAAGTATGTATGACTCCGCTTCAT 248
QY      84 ProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsn 103
Db      249 CCGACACAGGGGAGAGAGGTGGTCTGTATGGCCGCAATGTACAGCCAGGTGCCCATGAAC 308
QY      104 MetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrp 123
Db      309 ATGACCATCAGCTGGCTGCATGCTCACATTTACAGGAAGACCCCAACCGTGTCTTG 368
QY      124 GlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThr 143
Db      369 CAGTGGTGAATCAATCTTCAATGCAATGTTAACTACTCCAAACCGCAGTGTGACACT 428
QY      144 ProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAla 163
Db      429 CCCATCAGCTGTGAGGACAGCTGGGACAGCCATGTGAGTGCACACCTGAGCTGTGCC 488
QY      164 ThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheVal 183
Db      489 ACGCCCTGGGACTCAATCCCTCACCAGACACCTGCCCCCTTGGTGGCAGATTGTGTG 548
QY      184 ProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
Db      549 CCTTTTGCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGGACAGAGAG 608
QY      204 LeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThr 223
Db      609 CTGCAAGTGGGCATCCCGGTGCTGATGAGGCAAGGTCAAGGCTTGGTACTCGGTGACT 668
QY      224 AlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAla 243
Db      669 GCAGCCAAGCAGGAATCTTCCAGGTGGGATTTCAAGAACTGTGATGGCGATTCTGCC 728
QY      244 MetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db      729 ATGGCATCCACCACCTGATCATGACACTCTGAGAAAGAAAGACTTCTCTGAAG 782

```

RESULT 7
 BX355800 995 bp mRNA linear EST 05-MAY-2003
 LOCUS BX355800 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION clone CS0D1004YC06 5-PRIME, mRNA sequence.
 ACCESSION BX355800
 VERSION BX355800.1 GI:30384010

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 1 (bases 1 to 995) Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 BP 191 91006 Evry cedex - France
 Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1004BB03QPI&cluster=5549.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1004BB03QPI.

FEATURES
 source location/Qualifiers
 1..995
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1004YC06"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 226 a 281 c 281 g 193 t 14 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.15e-141 Length: 995
 Score: 1290.00 Matches: 252
 Percent Similarity: 96.55% Conservative: 0
 Best Local Similarity: 96.55% Mismatches: 9
 Query Match: 95.91% Indels: 0
 DB: 13 Gaps: 0

US-09-990-415A-2 (1-261) x BX355800 (1-995)

```

QY      1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
Db      91 ATGGAAGCAAAATGGGTGAATGCTTTAGACATCAACATCCAGGAACCTCGTGGAC 150
QY      21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db      151 CAAAGTACTTTCTGGGACAGAGCCCGGACACTTTTCACTGTACTGAWCCCGAAATCYG 210
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db      211 CYGCGTCCGGGGACAGCTGGAAGCTTCTCGGAACATCGTGACAGACMACAGGGCCGC 270
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db      271 GTGTGACCCCAAGGATCACCGAGGACCAAGTGTGAGGGCCAAAGTATGTATGACTCC 330
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db      331 GCCTTCCATCCGACACACAGGGGAGAGGKGGTCTGATTTGGCCGCAATGTCAAGGCTG 390
QY      101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db      391 CCCATGAACATGACCAACACTGGCTGCAATGCATTCATTCACATTCACATTCATTCATTC 450
QY      121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db      451 GTGTTTGGCAGTGGGTGAATCAGTCTTCAATGCAATTTGTTAACTACTCCAAACGCGAGT 510
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db      511 GGTGACACTCCCATCACTGTGAGGACAGTGGGACAGCCTATGTGAGTGCACCACTGGA 570
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180

```


Db 571 GCWGTGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCCTGCCCCCTTGTCGCGC 630
Qy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
|||||
Db 631 AGATTGTGCCCCCTTTCAGACAGTGGCAGCTGCCAAGTCGATCAACATCCCCCTGATGAGG 690
Qy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
|||||
Db 691 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGCAGGTACAGAGCCTTGCTAC 750
Qy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
|||||
Db 751 TCGGTGACTGCAGCCAGCAGGAATCTTCCAGGTGTGATTTCAAGAATCTGCATGGCG 810
Qy 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
|||||
Db 811 ATTCCTGCCATGGCCATCCCACTGATCATGACACTCTGGAGAAGAAAGACTTCCTG 870
Qy 261 Lys 261
|||
Db 871 AAG 873

RESULT 8
CB203758 898 bp mRNA linear EST 05-FEB-2003
LOCUS AGENCOURT_11288629 NIH_MGC_135 Mus musculus cDNA clone
DEFINITION IMAGE:30142456 5', mRNA sequence.
CB203758
VERSION CB203758.1 GI:28240427
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 898)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDA0051 row: 1 column: 17
High quality sequence stop: 676.

FEATURES

source

1. 898
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30142456"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRV; Site_2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5
, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'GACTAGTTCCTAGATCGCGAGCGCGCCCT(T)3' Tissue contributed by
, David Rowe. Library constructed by Resgen, Invitrogen
Corp."

BASE COUNT 190 a 255 c 253 g 199 t 1 others
ORIGIN

Alignment Scores:

6.64e-141 length: 898
Pred. No.: 1283.00 Matches: 248
Score:

Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 95.38% Mismatches: 5
Query Match: 95.39% Indels: 0
DB: 14 Gaps: 0
US-09-990-415a-2 (1-261) x CB203758 (1-898)

Qy 2 GluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpaspGln 21
|||||
Db 13 GAAAGCAAGATGGGTGACCTGCCCTTAATATCAACATCCAGAACTCGGTGGACCA 72
Qy 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 41
|||||
Db 73 AGCACATTCTAGGCAGAGCCCGCATTTCTTCACAGTCACATCCTCGAAATCTGCTG 132
Qy 42 LeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyVal 61
|||||
Db 133 CTGTCCGGGGAACAGCTGGAAGCTTCCCGGAACATCGTGCAGAAATTACAGGGCTGTGTG 192
Qy 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAla 81
|||||
Db 193 GCAACCCCGGGTCTCACTGAGACCAGCTATGGCGAGCCAAATACGTGTATGACTCAGCA 252
Qy 82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101
|||||
Db 253 TTTCATCCCGACACACGGGGAGAGGTGTCTTGTATGGCCGATGTACAGCCAGTGTGCC 312
Qy 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
|||||
Db 313 ATGACACATGACCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Qy 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSergly 141
|||||
Db 373 TTCTGGCAGTGGGTCAATCAGTCTTCAATGCTATGTGATTAATCTAATCGCAGCGGC 432
Qy 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
|||||
Db 433 GATGCTCCCATCACTGTGCAGCAGTTGGGACAGCCCTATGTGAGTGCCACCCTGGGGCT 492
Qy 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
|||||
Db 493 GTGGCTACTGCTCTGGGACTCAAGTCTCTCACCAGCAGCTGCCCTAGTCGCTGCA 552
Qy 182 PheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201
|||||
Db 553 TTCCGTGCCCTTTGACAGCTGTGGCGGCTGCCCACTGCATCAACATCCCCCTGATGAGCAG 612
Qy 202 ArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSer 221
|||||
Db 613 AGGAGAGCTGCAGGTGGGCATCCAGTGAATGATGAGCGCTGTCAGAGCCTTGCCACTTCG 672
Qy 222 ValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIle 241
|||||
Db 673 GTGACTGCTGCCAAGCAGGAATCTTCCAGGTGTGATATCAAGAATCGGAATGGCGATT 732
Qy 242 ProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
|||||
Db 733 CCCGCCATGGCCATTCGCCCGGTGATCATGAACACTCTGGAGAAGAAAGACTTCCTGAAG 792
RESULT 9
AK034514 2682 bp mRNA linear HTC 05-DEC-2002
LOCUS AK034514
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330209L13 product:sideroflexin 3, full
insert sequence.
AK034514
ACCESSION AK034514.1 GI:26330000
VERSION AK034514.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
3
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159
REFERENCE
AUTHORS

3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861
REFERENCE
AUTHORS
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Tomita,M.,
Quackenbush,J., Schriml,L.M., Staudli,F., Suzuki,R., Wagner,L.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamuya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Welter,C., Wittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohetsuki,S.
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
MEDLINE
PUBMED 11217851
REFERENCE
AUTHORS

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2682)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

COMMENT
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://phantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source
1. 2682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9330209L13"
/db_xref="taxon:10090"
/clone="9330209L13"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
36. 1001
/note="unnamed protein product; putative
sideroflexin 3 (MGDI/MT:2137679, GB|NM_053197, evidence:
BLASTN, 99%, match=2171)"
/codon_start=1
/protein_id="BAC28739.1"
/db_xref="GI:26330001"
/translation="MGDPLNINIQEPRMDQSTFLGRHFEVTDPNLLSGEQL
ASRNIVQNYRAGVATPGLTEDQLRAKVVYDSAFHPDGEKVLIGMSAQVPMNNTI
TGCMLTFYRKPTPTVFEWQVWQSFNAIVNYSNRSGDAPITVQQLGTAYSATGAVAT
ALGLSKLTKLPLPLVGRFVPEAAVAANCINIPLMRQRELVGIPPTDEAGORLGSV
TAAKQIGFQVVISRIGMATAMAIIPVIMNTLEKDFLKRPMGLAPLVGLVGRCLV
FATPLCALFPQRSSIHVTRLEPELRAQIQANPISIDVYYNKG.L"
2661. 2666
/note="putative"
2682
/note="putative"
/note="putative"
BASE COUNT 565 a 763 c 626 g 728 t
ORIGIN

polyA_signal
polyA_site
2682
Alignment Scores:
Pred. No.: 3.23e-140 Length: 2682
Score: 1283.00 Matches: 248
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 95.38% Mismatches: 5
Query Match: 95.39% Indels: 0
DB: 11 Gaps: 0

US-09-990-415a-2 (1-261) x AK034514 (1-2682)

OY 2 GUserlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTrpAspGln 21
|||||
DB 27 GAAGCAAGATGGGTGACCTGCCCTTAATATCAACATCCAGAACTCGGTGGACCAA 86
OY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 41
|||||
DB 87 AGCACATTTCTAGGCAGAGCCCGCATTTCTTACAGTCACTGATCCCCGAATCTGCTG 146
OY 42 LeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrrArgAlaGlyVal 61
|||||
DB 147 CTGTCCGGGGAACAGCTGGAAGCTTCCCGGAACATCTGCAAGATTACAGGGCTGTGTG 206
OY 62 ValThrProGlyIleThrGlnAspGlnLeuTrpArgAlaLeuTyrValTyrAspSerAla 81
|||||
DB 207 GCAACCCCGGCTCTCACTGAGACCAAGCTATGGCGAGCCAAATACGTATGACTCAGCA 266
OY 82 PheHisProAspThrGlyGluGlyValValLeuIleGlyArgMetSerAlaGlnValPro 101
|||||
DB 267 TTCCATCCGGACACAGGGGGAGAAAGTGTCTTGATTGGCCGTATGTCAGCCAGGTGCC 326

QY 102 MetasmetThrIleThrglyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
|||||
Db 327 ATGAACATGACCATTAATGCTGCTCATGCTCACCTTCTACAGAGAACTCCGACTGTGGTG 386
QY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSergly 141
|||||
Db 387 TTCTGGCAGTGGGTCAATCAGTCTCTCAATGCTATGTGATTAATTAATCGACGGCC 446
QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrglyAla 161
||| |||||
Db 447 GATGCTCCCATCAGTGTGAGCAGATTGGGACAGCCTATGTAGTGCCACCACTGGGGCT 506
QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
|||||
Db 507 GTGGCTACTGCTCTGGAGCTCAAGTCTCTCACCAAGCACTGCCCGCTAGTCGGTCCA 566
QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201
|||||
Db 567 TTGCTGCCCTTTGCAGCTGTGGCCGCTGCCAAGTGCATCAACATCCCCCTGATGAGGAG 626
QY 202 ArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSer 221
|||||
Db 627 AGGAGAGCTGCAGGTGGGCAATCCCAAGTACTGATGAGGCTGTGTAGAGGCTTGCCACTCG 686
QY 222 ValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIle 241
|||||
Db 687 GTGACTGCTGCCAAGAGGAGATCTTCCAGGTGTGATATCAAGAATCGAATGGCGATT 746
QY 242 ProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
|||||
Db 747 CCCGCCATGGCCATCCCGCGGTGATCATGAACACACTGTGAGAAGAAAGACTTCCTGAAG 806
RESULT 10
BQ959114 963 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_10031240 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482227
DEFINITION 5', mRNA sequence.
ACCESSION BQ959114
VERSION BQ959114.1 GI:22374592
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2664 row: n column: 20
High quality sequence stop: 635.
location/Qualifiers
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6482227"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by

ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 200 a 290 c 272 g 201 t
ORIGIN

Alignment Scores:
Pred. No.: 6.1e-133 Length: 963
Score: 1216.00 Matches: 237
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.41% Indels: 0
DB: 13 Gaps: 0

US-09-990-415a-2 (1-261) x BQ959114 (1-963)

QY 25 LeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSergly 44
|||||
Db 3 CTGGGCAAGAGCCCGGACATTTTTCACCTGTACTGATCTCTGAATCTGCTGTCGGG 62
QY 45 AlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrPro 64
|||||
Db 63 GCACAGCTGGAGCTTCTCGAACAATCGTGCAAGACTACAGAGGGCGGCTGTGACCCCA 122
QY 65 GlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisPro 84
|||||
Db 123 GGGATCACCGAGGACAGCTGTGAGGGGCCAAGTATGTATGACTCCGCCCTTCATCCG 182
QY 85 AspThrGlyGluLysValIleLeuIleGlyArgMetSerAlaGlnValProMetAsnMet 104
|||||
Db 183 GACACAGGGGAGAAGGTGTCTGATTTGGCCGATGTCAGCCCAAGTGCATGAACATG 242
QY 105 ThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGln 124
|||||
Db 243 ACCATCACTGGCTGCTGCATGCTTCTACAGGAAGACCCCAACCGTGTGTTCTGGCAG 302
QY 125 TrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerglyAspThrPro 144
|||||
Db 303 TGGGTGAATCAGTCTCTCAATGCCATTTGTTAACTACTCAACCCGAGTGTGACACTCCC 362
QY 145 IleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThr 164
|||||
Db 363 ATCACTGTGAGGACAGCTGGGACAGCCTATGTAGTGCCACCACTGGAGCTGTGGCCAG 422
QY 165 AlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro 184
|||||
Db 423 GCCCTGGGACTCAATATCCCTCACCAAGACACTGCCCCCTTGGTGGCAGATTGTGCC 482
QY 185 PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeu 204
|||||
Db 483 TTTCAGCAGAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGATGAGGACAGAGAGCTG 542
QY 205 GlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThrAla 224
|||||
Db 543 CAGGTGGGATCCCGGTGGTGTATGAGGACAGGTCAAGGCTTGCTACTCGGTGACTGCA 602
QY 225 AlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMet 244
|||||
Db 603 GCCAAGCAGGGAATCTTCACAGTGTGATTTCAAGAAATGTCATGGCGATTCTGCATG 662
QY 245 AlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
|||||
Db 663 GCCATCCCAACACATGATCAATGACACTCTGAGAGAAGAAAGACTTCCTGAAG 713
RESULT 11
BU506082 934 bp mRNA linear EST 12-SEP-2002
LOCUS AGENCOURT_10018523 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494119
DEFINITION 5', mRNA sequence.
ACCESSION BU506082
VERSION BU506082.1 GI:22812315
KEYWORDS EST.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 934)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM14049 row: n column: 08
High quality sequence stop: 651.
Location/Qualifiers
1..934
FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6494119"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 198 a 265 c 268 g 202 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.73e-132 length: 934
Score: 1212.00 Matches: 244
Percent Similarity: 96.18% Conservative: 8
Best Local Similarity: 93.13% Mismatches: 8
Query Match: 90.11% Indels: 3
DB: 13 Gaps: 0
US-09-990-415a-2 (1-261) x BU506082 (1-934)
QY 2 GluserlyMetGlygluLeuProLeuAspIleasnIleaglInluProArgTTPaspGln 21
Db 39 GAAAGCAAGATGGGTGACCTGCCCTTAATATACACATCCAGAACCTCGGTGGACCAA 98
QY 22 SerThrPheleuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 41
Db 99 AGCACATTCTAGGACAGAGCCCGGCATTCTTCACAGTCACTGATCCCCGAAATCTGCTG 158
QY 42 LeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTYrArgAlaGlyVal 61
Db 159 CTGTCCGGGGAACAGCTGGAACCTTCCCGAACATCGTGCAGAAATTACAGGGCTGCTG 218
QY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaIleValTYrValTyrAspSerAla 81
Db 219 GCAACCCCGGCTCTCACTGAGACCACTATGGCAGCCAAATACGTGTATGACTCAGCA 278
QY 82 PheHisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValPro 101
Db 279 TTCATCCGACACAGGGGAGAGGTGGTCTGTATGGCCGTATGTACGCCAGGTGCC 338
QY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
Db 339 ATGAACATGACCATTAATGCGCATGCTCACCCTCTACAGGAAGACTCCGACTGTGTG 398
QY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTYrSerAsnArgSerGly 141
Db 399 TTCGCGCAGTGGGTCAATCAATGCTTCAATGCTATGTGAATTACTCTAATCGCAGCGGC 458

QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
Db 459 GATGCTCCCATCACTGTGCAGCAGATTGGGAGACAGCCTATGTGATGCCACCACCTGGGCT 518
QY 162 ValAlaThrAlaLeuGlyLeuYsserLeuThrLysHisLeuProProLeuValGlyArg 181
Db 519 GTGGCTACTGCTCTGGAGCTCAAGTCTCTCACCAGCACCCTGCCCCGTAGTCGGTCA 578
QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201
Db 579 TTCGTGCCCTTTGCAGCTGTGGCCGCTGCCCACTGCATCAACATCCCCCTGATGAGGCAG 638
QY 202 ArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSer 221
Db 639 AGGAGCTGCAGGTGGGCATCCCACTGACTGATGAGGCTGTGACAGAGGCTTGCCACTCG 698
QY 222 ValThrAlaAlaLysGlnGlyIlePheGlnVal-ValIleSerArgIleCysMetAlaIle 241
Db 699 GTGACTGCTGCCAA-CAGGGAATCTTCCAGGTGGGTGATATCAAGAATCGGAATGGCGAT 757
QY 241 eProAlaMetAlaIle-ProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 261
Db 758 TCCCGCCATGCGCATTTCCCGCGGTGATCATGAACACTCTGGCAGAGANAGACTTCTGA 817
QY 261 ys 261
Db 818 AG 819
RESULT 12
BM548674 1040 bp mRNA linear EST 20-FEB-2002
LOCUS BM548674
DEFINITION AGENCOURT_6573430 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732220
5', mRNA sequence.
ACCESSION BM548674
VERSION BM548674.1 GI:18783439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1040)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM12733 row: p column: 13
High quality sequence start: 16
High quality sequence stop: 631.
Location/Qualifiers
1..1040
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5732220"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics

BASE COUNT 215 a 346 c 278 g 201 t
ORIGIN

Alignment Scores:

Pred. No.:	1.58e-130	Length:	1040
Score:	1196.00	Matches:	248
Percent Similarity:	96.89%	Conservative:	1
Best Local Similarity:	96.50%	Mismatches:	7
Query Match:	88.92%	Indels:	4
DB:	12	Gaps:	0

US-09-990-415A-2 (1-261) x BM548674 (1-1040)

```
QY      6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
         |||
Db      3 GGTGAATTGCTTTATACATCAACATCCAGGAACCTCGCTGGACCAAGT-ACCTTCCTG 61

QY      26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
         |||
Db      62 GGCAGAGCCCCGGCACATTTTCACTGTTACTGATGATCCTCGAAATCTGCTGTCGCGGGGCA 121

QY      46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
         |||
Db      122 CAGCTGAAGCTTCGCGAATCATCGTCAGAACTACAGGGCCGCGTGTGACCCCAAGG 181

QY      66 IleThrGluAspGlnLeuTrpArgAlaAlaTyrValTyrAspSerAlaPheHisProAsp 85
         |||
Db      182 ATCACCAGAGGACCACTGTGAGGGCCCAAGTATGTATGACTCCGCCCTTCATCCGGAC 241

QY      86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
         |||
Db      242 ACAGGGGAGAGGTGGTCTGATTGGCCGACATGCACGCCAGGTGCCCATGAACATGACC 301

QY      106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValAlaPheTrpGlnTrp 125
         |||
Db      302 ATCACTGGCTGCATGCTCACATTTCTACAGAGAACCCCAACCGTGTGTTGCGCAGTGG 361

QY      126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
         |||
Db      362 GTGAATCAGTCTTCAATGCCATGTGTAACACTCAACCGAGTGGTGACATCCCATC 421

QY      146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
         |||
Db      422 ACTGTAGGACGAGTGGGACAGACGCTATGTGAGTGCCACCACTGGAGCTGTGGCCACGGCC 481

QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValProPhe 185
         |||
Db      482 CTGGGACTCAAAATCCCTCACCAAGCACCTGCCCTTGCTCGGAGATTGTGCTCTTT 541

QY      186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
         |||
Db      542 GCAGCAGTGGCAGCTGCCAATGCATCAACATCCCTGATGAGGACAGAGAGCTGCAG 601

QY      206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
         |||
Db      602 GTGGGATCCCGGTGGCTGATGAGGACAGGTCAAGCCTTGCTACTCGGTGACTGCAGCC 661

QY      226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
         |||
Db      662 AG-CAGGGAATCTTCCAGGGGGGTGATTTC-AGAAATCTGCATGGCCATTCCTGGCCAGGCC 719

QY      246 IleProProLeuIleMetAspThrLeuGlu-LysLysAspPheLeuLys 261
         |||
Db      720 ATCCCCCCCCGTGATCATGACACTCTGGAAAAAATACTTCTTAAAA 768
```

RESULT 13
BG775086 961 bp mRNA linear EST 15-MAY-2001
LOCUS BG775086
DEFINITION 602650054F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760989 5',
 mRNA sequence.
ACCESSION BG775086
VERSION BG775086.1 GI:14045403

KEYWORDS
SOURCE
ORGANISM

EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-f@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1612 row: 1 column: 14
High quality sequence stop: 765.
Location/Qualifiers
1. .961
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4760989"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source

BASE COUNT 210 a 269 c 276 g 206 t
ORIGIN

Pred. No.:	9.45e-130	Length:	961
Score:	1189.00	Matches:	236
Percent Similarity:	99.16%	Conservative:	0
Best Local Similarity:	99.16%	Mismatches:	2
Query Match:	88.40%	Indels:	1
DB:	12	Gaps:	0

US-09-990-415A-2 (1-261) x BG775086 (1-961)

```
QY      24 PheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSer 43
         |||
Db      2 TTCCTGGCAGAGCCCGCACTTTTCACTGTACTGATCCTCGAAATCTGCTGTGCC 61

QY      44 GlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThr 63
         |||
Db      62 GGGGCACAGCTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGGCCGCGTGTGACC 121

QY      64 ProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHis 83
         |||
Db      122 CCAGGATCACCGAGACCACTGTGAGAGGCCCAAGTATGTATGACTCCGCCCTTCAT 181

QY      84 ProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsn 103
         |||
Db      182 CCGGACACAGGGGAGAGGTGGTCTGATTTGGCCGATGTACAGCCAGGTGCCCATGAAC 241

QY      104 MetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrp 123
         |||
Db      242 ATGACCATCACTGGCTGCATGCTCACATTTACAGGAAGACCCCAACCGTGTGTTCTGG 301

QY      124 GlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThr 143
         |||
Db      302 CAGTGGTGAATCAGTCCCTTCAATGACCATTTGTTAACTACTCCAAACGACAGTGTGACACT 361
```

QY 144 ProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAla 163
 Db 362 CCCATCACTGTGAGCAGCTGGGGACGCTATGTGAGTCCACCACTGGAGCTGTGGCC 421

QY 164 ThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheVal 183
 Db 422 ACGGCCCTGGGACTCAAAATCCCTCACCAGCACCTGCCCCCTGTGCGCAGATTGTGT 481

QY 184 ProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
 Db 482 CCTTTGACAGCAGTGGCAGCTGCCAATGCATCAACATCCCCCTGATGAGCAGAGAGAG 541

QY 204 LeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThr 223
 Db 542 CTGCAGGTGGGCATCCCGGTGGCTGATGAGCAGGTGTCAGAGGCTTGCTACTCGGTGACT 601

QY 224 AlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAla 243
 Db 602 GCAGCCACGACGAGGAATCTTCCAGGTGTGATTTCAGAAGATCTGCATGGCGATTCTGCC 661

QY 244 MetAlaIleProProLeuIleMetAspThrLeuGlyLysLysAspPheLeuLys 261
 Db 662 ATGACATCCCAACCATG-ATCATGGACACTCTGGAGAGAAAGACTTCTCTGAAG 714

RESULT 14
 BQ425697 862 bp mRNA linear EST 23-MAY-2002
 LOCUS AGENCOURT_7887371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6166944
 DEFINITION 5', mRNA sequence.

ACCESSION BQ425697
 VERSION BQ425697.1 GI:21121012
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 862)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://limage.llnl.gov
 Plate: LLM13528 row: b column: 01
 High quality sequence stop: 603.

FEATURES
 source
 1..862
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6166944"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 200 a 249 c 246 g 166 t 1 others
 ORIGIN:

Alignment Scores:
 Pred. No.: 9.35e-129 Length: 862
 Score: 1180.00 Matches: 234
 Percent Similarity: 94.82% Conservative: 4
 Best Local Similarity: 93.23% Mismatches: 10

Query Match: 87.73% Indels: 3
 DB: 13 Gaps: 2
 US-09-990-415a-2 (1-261) x BQ425697 (1-862)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 Db 56 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGCTGGGAC 115

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 Db 116 CAAAGTACTTCTCGGGCAGAGCCCGGCACTTTTTCACGTGTACTGATCTCGAAATCTG 175

QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 Db 176 CTGCTGTCCGGGGCACACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCCGGC 235

QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80
 Db 236 GTGTGTACCCCAAGGATCACCGAGGACCACTGTGAGGGCCAAATATGTATGACTCC 295

QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 Db 296 GCCTTCCATCCGACACACAGGGAGAGAGGTCTCTGATTGGCCGCATGCACGCCAGGTG 355

QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 Db 356 CCCATGAACATGACCATCCTGCTGCTGCATCTACATCTACAGAGAACCCCAACCGTG 415

QY 121 ValPheTyrGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 Db 416 GTGTCTCGGACGTGGGTGAATCAGTCTCTCAATGCCATGTGTAACACTCCCAACCGCAGT 475

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThr-ThrG1 160
 Db 476 GGTGACACTCCCATCCTGAGGACAGTGGGGACAGACCCATGTGAGTCCACCCACTGG 535

QY 160 ValAlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValG1 180
 Db 536 AGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGGCCCTTGGTGG 595

QY 180 YArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetAr 200
 Db 596 CAGATTGTGCCCTTTTGCAGCAGTGGCAGTGCACCACTGCATCAACATCCCTGTATGAA 655

QY 200 GglnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTy 220
 Db 656 GGAGAGAGAGCTGCANGTGGGCATCCCGGTGGCTGATGAGCAGAGTCAACAAGCTTGCTA 715

QY 220 rSerValThrAlaAlaLysGlnGlyIlePheGln---ValValIleSerArgIleCysMe 239
 Db 716 CTCGGTACTGTCAGCCAGCAAGCAGGAATCTTCCAGGGGGTGGATTTCAGAAGATCTGCCA 775

QY 239 tAla---IleProAlaMetAlaIleProPro 248
 Db 776 GGGGCGAATCCCGGCCCTGGCCCATCCCCC 806

RESULT 15
 CA317398 731 bp mRNA linear EST 26-NOV-2002
 LOCUS UI-M-FW0-cdn-g-19-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
 DEFINITION IMAGE: 6811916 5', mRNA sequence.
 ACCESSION CA317398
 VERSION CA317398.1 GI:24535522
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 731)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source
Seq primer: PYX-5.
Location/Qualifiers
1..731

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6811916"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FW0"

/note="Organ: Brain; Vector: PYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemlin Chlin, Ph.D.,
program coordinator."

BASE COUNT 163 a 201 c 206 g 159 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 3.33e-127 length: 731
Score: 1166.00 Matches: 228
Percent Similarity: 97.08% Conservative: 5
Best Local Similarity: 95.00% Mismatches: 6
Query Match: 86.69% Indels: 1
DB: 14 Gaps: 0

US-09-990-415A-2 (1-261) x CA317398 (1-731)

OY 2 GluserlysmetGlyGluLeuProleuAspIleasnileglnGluProargTrpaspGln 21
|||||
Db 12 GAAAGCAAGATGGGTGACCTGCCCCCTTAATATCAACATCCAGAACCTCGGTGGACCAA 71
OY 22 SerThrpheluGlyArgAlaArgHisphetherValThrAspProArgasnLeuLeu 41
|||||
Db 72 AGCACATTCTAGGCGAGAGCCCGGCACTTCTTCACAGTCACCTGATCCCCGAAATCTGCTG 131
OY 42 LeuSerGlyAlaGlnLeuGluAlaSerArgasnIleValGlnAsnTyrArgAlaGlyVal 61
|||||
Db 132 CTGTCCGGGGAGACAGCTGGAAAGCTTCCCGGAACATCGTGACAGATTACAGGGCTGTGTG 191
OY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaIleTyrValTyrAspSerAla 81
|||||
Db 192 GCAACCCCGGCTCTCACTGAGGACAGCTATGGCGAGCCAAATACGTATGACTCAGCA 251
OY 82 PheHisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValPro 101
|||||
Db 252 TTCCATCCGACACGGGAGAGAGAGTGTGTGATGGCCGTATGTCAAGCCAGGTGCC 311
OY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
|||||

Db 312 ATGACATGACCATTTACTGGCTGCATGCTCACCTTCTACAGAGAACTCCGACTGTGCTG 371
OY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
|||||
Db 372 TTCTGGCAGTGGGTCAATCAGTCTTCAATGCTATTGTGAATTAATCTAATCGCAGCGCG 431
OY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
|||||
Db 432 GATGCTCCATCACTGTGCAGCAGTGGGGACAGCCATATGTGAGTGCACCACTGGGGCT 491
OY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArg 181
|||||
Db 492 GTGGCTACTGCTGTGGAGCTCAAGTCTCTCACCAAGCACCTGCCCTAGTCGTCGA 551
OY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeu-MetArgG1 201
Db 552 TTCTGCCCCCTTTGCAGCTGTGGCCGCTGCCAAGTGCATCAACATCCCCCTNGATGAGCGA 611
OY 201 nArgGlnLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSe 221
Db 612 GAGGAGCTGCAGAGTGGGCATCCAGTCACTGATGAGGCTGTGCAGAGGCTTGCCACTC 671
OY 221 rValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetaLa 240
Db 672 GTGACTGCTGCCANACAGGGAATCTTCCAGGTGGTGTATATCAAGAAATCGGAATGCGG 729

Search completed: October 16, 2003, 17:34:28
Job time : 2083 secs


```

OY 349 GCTGTCGGCGTCACGCGTACGCTCCCGCTGATGGCTGGAGGCGCGCGGCGACAGC 408
Db 146 GCTGTCGGCGTACGCGTACGCTCCCGCTGATGGCTGGAGGCGCGCGGCGACAGC 205
OY 409 GGAGCGAGAGAGAGAGGCGGCTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 468
Db 206 GGAGCGAGAGAGAGAGGCGGCTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTG 265
OY 469 AATTGCTTTAGACATCAACATCCAGAGAACCTCGCTGGAGCAAAAGTACTTCTGGGCA 528
Db 266 AATTGCTTTAGACATCAACATCCAGAGAACCTCGCTGGAGCAAAAGTACTTCTGGGCA 325
OY 529 GAGCGCGGCGCTTTTCACTGTACTGATCCTCGAAATCTGCTGCTCGGCGGCGACAGC 588
Db 326 GAGCGCGGCGCTTTTCACTGTACTGATCCTCGAAATCTGCTGCTCGGCGGCGACAGC 385
OY 589 TGAAGCTTCTCGAATCCTGTCAGAACTACAGGCGCGGCTGTGAGCGCGGCGATCA 648
Db 386 TGAAGCTTCTCGAATCCTGTCAGAACTACAGGCGCGGCTGTGAGCGCGGCGATCA 445
OY 649 CCGAGGACGAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCGAGACAG 708
Db 446 CCGAGGACGAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCGAGACAG 505
OY 709 GGGAGAGGTGCTCTGATTTGGCGCATGTCAGCCGAGTCCCATGACATGACCATCA 768
Db 506 GGGAGAGGTGCTCTGATTTGGCGCATGTCAGCCGAGTCCCATGACATGACCATCA 565
OY 769 CTGCTGATGCTCATTCTACAGAGAGAGCCGAGTGTGCTGAGAGTGGGTGA 828
Db 566 CTGCTGATGCTCATTCTACAGAGAGAGCCGAGTGTGCTGAGAGTGGGTGA 625
OY 829 ATCAGTCTTCAATGCCATGTGTAATCTACTCCAAACCGAGTGTGACACTCCCATCACTG 888
Db 626 ATCAGTCTTCAATGCCATGTGTAATCTACTCCAAACCGAGTGTGACACTCCCATCACTG 685
OY 889 TGAGGCGAGCTGGGAGAGAGCTATGTAGTGCACACCTGAGCTGTGGCCAGCGGCTGG 948
Db 686 TGAGGCGAGCTGGGAGAGAGCTATGTAGTGCACACCTGAGCTGTGGCCAGCGGCTGG 745
OY 949 GACTCAATCCCTCACAAGCAACCTGCGCGCTGTCGAGATTTGTGCTTTCAG 1008
Db 746 GACTCAATCCCTCACAAGCAACCTGCGCGCTGTCGAGATTTGTGCTTTCAG 805
OY 1009 CAGTGGAGCTGCCAATCTCATCAACATCCCTGATGAGGAGAGAGAGCTGCAAGTGG 1068
Db 806 CAGTGGAGCTGCCAATCTCATCAACATCCCTGATGAGGAGAGAGAGCTGCAAGTGG 865
OY 1069 GCATCCCGGCTGTATGAGGAGAGTCAAGGCTTGGCTACTCGGTGACTGCAAGCAAGC 1128
Db 866 GCATCCCGGCTGTATGAGGAGAGTCAAGGCTTGGCTACTCGGTGACTGCAAGCAAGC 925
OY 1129 AGGGAATCTTCCAGGCTGTGATTTCAAGATCTGCATGGCATTCCTGCGCATGCC 1188
Db 926 AGGGAATCTTCCAGGCTGTGATTTCAAGATCTGCATGGCATTCCTGCGCATGCC 985
OY 1189 CACCACTGATCATGACACTCTGGAGAAAGAAAGACTTCTGAG 1232
Db 986 CACCACTGATCATGACACTCTGGAGAAAGAAAGACTTCTGAG 1029

```

RESULT 2

```

US-08-905-223-165
; Sequence 165, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear

```

```

STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 165:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: s1g-peptide
LOCATION: 70..108
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 3.5
US-08-905-223-165
Query Match 5.38; Score 65.8; DB 3; Length 315;
Best Local Similarity 56.18; Pred. No. 1.2e-08;
Matches 124; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
OY 496 AACCTGCTGGACCAAGTACTTCTCGGAGAGAGCGGCGACTTTTCACTGTACTG 555
Db 94 AACCCGCTTCCAGCAGACGCTCTCTATGCGCGCTTCAGGCACTTCTGATATCATCG 153
OY 556 ATCTCGAATCTGCTGCTGTCGCGGCGAGAGCTGGAAGCTTCTCGAATCATCGTCA 615
Db 154 ACCCTGCGACACTTCTTGTCACTGAGAGAGCTCTCAGAGAGGCTGTGACGCTGAGG 213
OY 616 ACTACAGGCGCGCGTGTGACCCAGGAGATCACGAGAGACGCTGTGAGGCGCAAGT 675
Db 214 ACTATTAAGCATGAGGAGCCCTGCGCGCGGCGGCTCAGCAATGAGACGCTGTGAGTGCACAGA 273
OY 676 ATGTATGACTCGCGCTTCCATCCGAGACACAGGGAGAG 716
Db 274 AATCAAGCAGGCTATTTCACTCCGAGACCAATGAGAG 314

```

RESULT 3

```

US-09-922-146-3
; Sequence 3, Application US/09922146
; Patent No. 6566133
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
; FILE REFERENCE: RTS-0252
; CURRENT APPLICATION NUMBER: US/09/922,146
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 16:00:17 ; Search time 77 Seconds
(without alignments)

1496.117 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPLDINIQEPRWD.....PAMAIPLIMDTLEKKDFLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09990415/runat_16102003_073807_9464/app_query.fasta_1.455
-DB-Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415 -CGN_1_1_85 -runat_16102003_073807_9464 -NCPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1729	4	US-09-620-312D-1016 Sequence 1016, Ap
2	179	13.3	315	3	US-08-905-223-165 Sequence 165, App
3	96	7.1	2943	1	US-07-923-976-3 Sequence 3, Appl1
4	92	6.8	2563	6	5422248-1 Patent No. 5422248
5	89.5	6.7	3024	1	US-07-923-976-7 Sequence 7, Appl1
6	89	6.6	1968	4	US-09-134-001C-2694 Sequence 2694, Ap
7	88	6.5	669	4	US-09-252-991A-1852 Sequence 1852, Ap
8	88	6.5	1185	4	US-09-252-991A-1931 Sequence 1931, Ap
9	85	6.3	552	4	US-09-252-991A-2024 Sequence 2024, Ap
10	84.5	6.3	2855	1	US-07-923-976-5 Sequence 5, Appl1
11	82	6.1	50341	1	US-08-247-901C-1 Sequence 1, Appl1
12	82	6.1	50341	2	US-09-075-904-1 Sequence 1, Appl1

13	82	6.1	52297	3	US-09-426-436-1	Sequence 1, Appl1
14	82	6.1	52297	4	US-08-705-557-1	Sequence 1, Appl1
15	81	6.0	663	4	US-09-252-991A-9294	Sequence 9294, Ap
16	81	6.0	918	4	US-09-252-991A-9303	Sequence 9303, Ap
17	81	6.0	1428	4	US-09-252-991A-9266	Sequence 9266, Ap
18	81	6.0	1545	4	US-09-252-991A-9274	Sequence 9274, Ap
19	81	6.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
20	81	6.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
21	80.5	6.0	1153	4	US-09-372-448A-5	Sequence 5, Appl1
22	80	5.9	1866	4	US-09-252-991A-3178	Sequence 3178, Ap
23	79.5	5.9	1884	4	US-09-687-538B-5	Sequence 5, Appl1
24	79.5	5.9	2237	1	US-08-463-620-1	Sequence 1, Appl1
25	79.5	5.9	2237	2	US-08-224-917-1	Sequence 1, Appl1
26	79.5	5.9	2237	2	US-08-914-853-1	Sequence 1, Appl1
27	79.5	5.9	2237	5	PCT-US95-03934A-1	Sequence 1, Appl1
28	79.5	5.9	12412	1	US-08-390-878-18	Sequence 18, Appl1
29	79.5	5.9	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
30	79	5.9	1419	4	US-09-252-991A-12594	Sequence 12594, A
31	79	5.9	2238	4	US-09-252-991A-13219	Sequence 13219, A
32	78.5	5.8	2340	4	US-09-786-240-29	Sequence 29, Appl1
33	78.5	5.8	2350	4	US-09-620-312D-1094	Sequence 1094, Ap
34	78.5	5.8	9603	4	US-09-203-895-1	Sequence 1, Appl1
35	78.5	5.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
36	78	5.8	2250	4	US-09-221-017B-1033	Sequence 1033, Ap
37	77.5	5.8	2028	4	US-09-221-017B-1001	Sequence 1001, Ap
38	77.5	5.8	6027	4	US-09-620-312D-517	Sequence 517, Ap
39	77.5	5.8	25165	4	US-09-453-702B-39	Sequence 39, Appl1
40	77	5.7	864	4	US-09-252-991A-1297	Sequence 1297, Ap
41	77	5.7	1050	4	US-09-252-991A-1269	Sequence 1269, Ap
42	77	5.7	1308	4	US-09-252-991A-1179	Sequence 1179, Ap
43	77	5.7	1878	3	US-08-996-139-14	Sequence 14, Appl1
44	77	5.7	1878	3	US-08-995-659-14	Sequence 14, Appl1
45	77	5.7	1878	3	US-09-215-649A-14	Sequence 14, Appl1

ALIGNMENTS

RESULT 1
US-09-620-312D-1016
; Sequence 1016, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (247)..(1371)
US-09-620-312D-1016

Alignment Scores:

Pred. No.:	5.72e-165	Length:	1729
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-990-415A-2 (1-261) x US-09-620-312D-1016 (1-1729)

```
OY 1.MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIuProArgTrpasp 20
DB 247.ATGGAAGCAAAATGGGTGAATTGCCCTTTAGACATCAACATCCAGGAACCTCGCTGGGAC 306
OY 21.GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 307.CAAAGTACTTCTCTGGGCAGAGACCCCGCACTTTTTCACGTGTACTGTATCTCGAAATCTG 366
OY 41.LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 367.CTGCTGTCCGGGGCAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCCGGC 426
OY 61.ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
DB 427.GTGTGACCCCGAGGATCACCGAGACCAAGCTGTGGAGGGCCAAATATGTATGACTCC 486
OY 81.AlaPheHisProAspThrGlyGlyLysValIleGlyArgMetSerAlaGlnVal 100
DB 487.GCCTTCCATCCGGACACAGGGGAGAAGTGTCTGTATGGCCGATGTACAGCCAGGTG 546
OY 101.PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 547.CCCATGAACATGACCATCTGCTGCTGATGCTACATCTTACAGGAAGACCCCAACCGTG 606
OY 121.ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 607.GTGTCTTGGCAGTGGGTGAATCAGTCTTCAATGCCATGTGTAACTACTCCAACCGCAGT 666
OY 141.GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 667.GGTGACACTCCCATCTGCTGAGGCAAGCTGGGACAGCCCTATGTGAGTGGCCACCACTGGA 726
OY 161.AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 727.GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACTGCCCTTGTGCGGC 786
OY 181.ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 787.AGATTGTGCCCCCTTTCAGCAGCTGCGCACTGCCAATGCATCAACATCCCTGATGAGG 846
OY 201.GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 847.CAGAGAGAGCTGCAGGTGGGCACTCCCGGTGCTGATGAGGCAGGTGCAGAGCTTGCTAC 906
OY 221.SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 907.TCGGTGACTGCAGCCAAAGCAGGGAATCTTCCAGGTGTGATTTCAAGAACTCTGCATGGCG 966
OY 241.IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
DB 967.ATTCTGCGCATGGCCATCCACCACTGATGACACTCTGTGAGAGAAGAAGACTTCTCTG 1026
OY 261.Lys 261
DB 1027.AAG 1029
```

RESULT 2

US-08-905-223-165

Sequence 165, Application US/08905223

Patent No. 6222029

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duclert, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 165:

SEQUENCE CHARACTERISTICS:

LENGTH: 315 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: DOUBLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Homo Sapiens

TISSUE TYPE: Brain

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 70..108

IDENTIFICATION METHOD: Von Heijne matrix

OTHER INFORMATION: score 3.5

OTHER INFORMATION: seq MHLISNWPASS/RR

US-08-905-223-165

Alignment Scores:

Pred. No.:	2.62e-14	Length:	315
Score:	179.00	Matches:	29
Percent Similarity:	66.22%	Conservative:	20
Best Local Similarity:	39.19%	Mismatches:	25
Query Match:	13.31%	Indels:	0
DB:	3	Gaps:	0

US-09-990-415A-2 (1-261) x US-08-905-223-165 (1-315)

```
OY 16.GluProArgTrpAspGlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThr 35
DB 93.AAACCCTGCTCCAGCAGACGCTCTTATGGCCGCTTCAGGCACCTCTTGATATCATC 152
OY 36.AspProArgAsnLeuLeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGln 55
DB 153.GACCCCTGCACACTCTTGTCTACTGAGAGACGCTCTCAGAGAGGCTGTGACGCTGGAG 212
OY 56.AsnTyrArgAlaGlyValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLys 75
DB 213.GACTATAAGCATGGGACCCCTGCGCCGGGGGTCAACCAATGAACAGCTCTGAGTGCACAG 272
OY 76.TyrValTyrAspSerAlaPheHisProAspThrGlyGlyLys 89
```

```

Db      273 AAAATCAGACGCTATTCTACATCCGACACCAATGAGAAG 314
      RESULT 3
US-07-923-976-3
; Sequence 3, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
;   APPLICANT: Nagata, Shigekazu
;   APPLICANT: Fukunaga, Rikio
;   TITLE OF INVENTION: DNA Encoding Granulocyte
;   TITLE OF INVENTION: Colony-Stimulating Factor Receptor
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Jones, Tullar & Cooper, P.C.
;     STREET: P.O. Box 2266 Eads Station
;     CITY: Arlington
;     STATE: Virginia
;     ZIP: 22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentia Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/07/923,976
;     FILING DATE: 19920922
;     CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 74539/1990
;       FILING DATE: 23-MAR-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 176629/1990
;       FILING DATE: 03-JUL-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: PCT/JP91/00375
;       FILING DATE: 22-MAR-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Hellwege, James W.
;       REGISTRATION NUMBER: 28,808
;       REFERENCE/DOCKET NUMBER: 514853
;       TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 703-415-1500
;       TELEFAX: 703-415-1508
;     INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 2943 base pairs
;         TYPE: NUCLEIC ACID
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;       MOLECULE TYPE: cDNA to mRNA
;       FEATURE:
;         NAME/KEY: CDS
;         LOCATION: 170..2677
;       US-07-923-976-3

Alignment Scores:
Pred. No.:      0.0693      Length:      2943
Score:          96.00      Matches:       59
Percent Similarity: 42.86%  Conservative: 34
Best Local Similarity: 27.19%  Mismatches:  89
Query Match:    7.14%      Indels:      36
DB:             1         Gaps:         12

US-09-990-415A-2 (1-261) x US-07-923-976-3 (1-2943)

QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLys-----TyrValTyr 78
      ::::::::::: ::::: ::::: ::::: :::::
Db      1676 ATCGTAGCTCCC--TTGTACCAGACACCATGGGACCCCTCCAGCATGTCTATGCTTAC 1732

QY      79 AspSerAlaPheHisProAspThrGlyGluLysValValLeu-----IleGlyArgMet 96
      ::::: ::::: ::::: ::::: :::::
Db      1733 TCTCAAGAAATGGCTCCCTCCCATGCCCCAGACAGCTGCTAAAGCACATGCGACAAGACC 1792

```

```

QY      97 SerAlaGln-----ValPrometAsnMetThrIleThrGlyCysMetLeuThrPhe 113
      ::::: ::::: ::::: :::::
Db      1793 TGGGCACAGCTGAGTGGGTGCTGAGCCCCCTGAGCTGGGGAAGAGCCCTTACCAC 1852

QY      114 TyrArgLysThrProThrValValPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
      ::::: ::::: ::::: :::::
Db      1853 TAC-----ACCATCTCTGGACCAACGCTCAGAACCAAGTCTCTCCGCC 1897

QY      133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThr----- 146
      ::::: ::::: ::::: :::::
Db      1898 ATCCTGAATGCCCTCTCCCTGCTGCTTGTCTCATGGCCCTGAGCCCGCCAGTCTGTAT 1957

QY      147 ---ValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
      ::::: ::::: ::::: :::::
Db      1958 CACATCCACCTCATGGCTGCCAGCCAGGCTGGGGCCACCAACAGTACAGTCCACCCCTG 2017

QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValPro--- 184
      ::::: ::::: ::::: :::::
Db      2018 ATGACCTTGACCCAGAGGGGTGCGAGCTACACATCATCTGGCCCTGTCGG--CCTCCT 2076

QY      185 ---PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
      ::::: ::::: ::::: :::::
Db      2077 GCTGTGCTCACCCTGCTCTGTGGAAGTGCCTGCTCTG---TTGCAGCCCCACACAGAA 2133

QY      204 LeuGln-ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerVal-- 222
      ::::: ::::: ::::: :::::
Db      2134 GAATCCCTCTGGCCCAAGTGTCCAGACCCAGCTCACAGCAGCTGGGCTCTGGGTGCC 2193

QY      223 -ThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIlePr 242
      ::::: ::::: ::::: :::::
Db      2194 CACAATCATGAGAGGATGCTTCAG-----CTGCC 2226

QY      242 oAlaMetAlaIleProLeuIleMetAspThrLeuGluLysAsp 258
      ::::: ::::: ::::: :::::
Db      2227 CGGCTTGGCAGCCACCATCACCAAGCTCAGAGTGTGAGAGAGAT 2275

RESULT 4
5422248-1
; Patent No. 5422248
; APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
; BENSON, CURTIS M.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
; STIMULATING FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/6,183
;   FILING DATE: 15-JAN-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 587,329
;     FILING DATE: 24-SEP-1990
;     APPLICATION NUMBER: 522,952
;     FILING DATE: 03-APR-1990
;     APPLICATION NUMBER: 416,306
;     FILING DATE: 03-OCT-1989
;     APPLICATION NUMBER: 412,816
;     FILING DATE: 26-SEP-1989
;   SEQ ID NO: 1:
;     LENGTH: 2503
;     5422248-1

Alignment Scores:
Pred. No.:      0.183      Length:      2563
Score:          92.00      Matches:       58
Percent Similarity: 42.86%  Conservative: 35
Best Local Similarity: 26.73%  Mismatches:  89
Query Match:    6.84%      Indels:      36
DB:             6         Gaps:         12

US-09-990-415A-2 (1-261) x 5422248-1 (1-2563)

QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLys-----TyrValTyr 78
      ::::::::::: ::::: ::::: ::::: :::::

```



```

Db      1688 ATCGTGAAGTCCC--TTGTACAGAGACACCATGGAGCCCTCCAGCATGTCTATGCTTAC 1744
QY      79  AspSerAlaPheHisProAspThrGlyGluValValLeu-----IleGlyArgMet 96
Db      1745 TCTCAAGAAATGGCTCCCTCCATGCCCCAGAGCTGCATCTAAAGCATATGGCAAGACC 1804
QY      97  SerAlaGln-----ValProMetAsnMetThrIleThrGlyCysMetLeuThrPhe 113
Db      1805 TGGGCACAGCTGGAGTGGTGGCTGAGACCCCTGAGCTGGGAGAGAGCCCTTACCCAC 1864
QY      114 TyrArgLysThrProThrValAlaPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
Db      1865 TAC-----ACCATCTTCTGGACCAACGCTCAGAACCAAGTCTCTCTCGCC 1909
QY      133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThr----- 146
Db      1910 ATCCTGAATGCTCCTCCCTGCTTGTCTTCTCCATGCTGAGAGCCCGCAGTCTGTAT 1969
QY      147 ---ValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db      1970 CACATCCACCTCATGCTGCCAGCCAGGCTGGGGCCACCAAGTACAGTCTCCTACCCCTG 2029
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro--- 184
Db      2030 ATGACCTTGACCCAGAGGGGTCCGAGCTACACATCATCTGGCCCTGTCCG-CTTCT 2088
QY      185 ---PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
Db      2089 GCTGTGCTCACCCTGCTCTGTGGAAGTGCCTGGCTGTG---TTGACGGGGGAACAGGAA 2145
QY      204 LeuGln-ValGlyIleProValAlaAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerVal-- 222
Db      2146 GAATCCCCCTTGCGCCAGTTGCCCCAGACCCAGCTCACAGCAGCCTGGCTCCTGGGTGCC 2205
QY      223 -ThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIlePr 242
Db      2206 CACAATCATGGAGAGGATGCTTCCAG-----CTGCC 2238
QY      242 AlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAsp 258
Db      2239 CGGCTTGCGACGCCACCATCACCAAGCTCACAGTGTGAGAGAGAT 2287

```

```

RESULT 5
US-07-923-976-7
; Sequence 7, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte
; TITLE OF INVENTION: Colony-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923, 976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3024 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 170..2758
; US-07-923-976-7

```

```

Alignment Scores:
Pred. No.: 0.509 Length: 3024
Score: 89.50 Matches: 55
Percent Similarity: 36.73% Conservative: 35
Best Local Similarity: 22.45% Mismatches: 90
Query Match: 6.65% Indels: 65
DB: 1 Gaps: 13

```

US-09-990-415a-2 (1-261) x US-07-923-976-7 (1-3024)

```

QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLys-----TyrValTyr 78
Db      1676 ATCGTGAAGTCCC--TTGTACAGAGACACCATGGAGCCCTCCAGCATGTCTATGCTTAC 1732
QY      79  AspSerAlaPheHisProAspThrGlyGluValValLeu-----IleGlyArgMet 96
Db      1733 TCTCAAGAAATGGCTCCCTCCATGCCCCAGAGCTGCATCTAAAGCATTTGGCAAGACC 1792
QY      97  SerAlaGln-----ValProMetAsnMetThrIleThrGlyCysMetLeuThrPhe 113
Db      1793 TGGGCACAGCTGGAGTGGGTGCTGAGCCCCCTGAGCTGGGGAAGAGCCCTTACCCAC 1852
QY      114 TyrArgLysThrProThrValAlaPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
Db      1853 TAC-----ACCATCTTCTGGACCAACGCTCAGAACAGTCTCTTCTCGCC 1897
QY      133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThr----- 146
Db      1898 ATCCTGAATGCTCCTCCCTGCTTGTCTCCATAGCCCTGAGAGCCCGCAGTCTGTAT 1957
QY      147 ---ValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db      1958 CACATCCACCTCATGGCTGCCAGCCAGGCTGGGGCCACCAACAGTACAGTCTCCTACCCCTG 2017
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db      2018 ATGACCTTGACCCAGAGGGGTCCGAGCTACACATCATCTGGGCTG-----TTC 2068
QY      186 AlaAlaValAlaAlaAlaAsnCysIle----- 194
Db      2069 GGCCTCTGCTGTGCTCACCCTGCTGTGGAAGTGCCTGGCTCTGTTCAGAGCCCAAC 2128
QY      195 -----AsnIleProLeuMetArgGlnArgGlnValGly 207
Db      2129 AGGAAGATCCCTCTGGCCCAAGTGTCCAGACCCAGCTCACAGCAGCCTGGGCTCTGG 2188
QY      208 IleProValAlaAspGlu-----AlaGlyGlnArgLeuGlyTyrSer 221
Db      2189 GTGCCACAAATCATGAGGAGCTGCCCGGAGCCAGACAGAGACAGTGGCTGGGCAAGACA 2248
QY      222 ValThrAlaAlaLys-----GlnGlyIlePheGlnValVal 233

```

```

Db      2249 TCTGAATGAGCCGTCTCTCACCACCATCCTTGTGTGACGATGCCCTTCAG----- 2302
               :::::  |||  |||||
Qy      234  Ileserargilecysmetalaleproalamealaleproleuilemetaspthr 253
               ::|||  ::  |||||::  |||
Db      2303 -----CTGCCCGGCTTGGACAGCCACCATCACCAAGCTCACA 2341
Qy      254  leuglulyslyasp 258
               ::  :::::|||
Db      2342 GTGCTGAGAGGAT 2356

RESULT 6
US-09-134-001C-2694
; Sequence 2694, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2694
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2694

Alignment Scores:
Pred. No.:      0.29      Length:      1968
Score:          89.00     Matches:      54
Percent Similarity: 33.33%  Conservative: 32
Best Local Similarity: 20.93%  Mismatches: 78
Query Match:    6.62%      Indels:      94
DB:             4         Gaps:      10

US-09-990-415a-2 (1-261) x US-09-134-001C-2694 (1-1968)
Qy      53  llevalglnasnlyrrargalaglyvalvalthrproglyilethrgluaspglnleutrp 72
               |||:::  ::  |||||  ::|||  |||||
Db      1219 ATTATGTGCTGTTCTTACGAGGTATTACTCAAGAAAT----- 1260
Qy      73  ArgAlalystyrvaltyraspserralaphehisproaspthrglylulysvalvalleu 92
               |||||:::  |||  |||  |||
Db      1261 -----AAATATATTACACGTAATTCGCCACAAGCAATTGAAGGCTTGAACCTACTTTA 1314
Qy      93  ileglyargmetseralaglnvalprometasmethrilethrglycysmethleuthr 112
               |||  ::  :::::|||||  :::::
Db      1315 ATCTATCCATTA-----CTTCAGTTTCATTTACTGGATTGTTAATGGTG 1359
Qy      113 pheTyrArglysthrProthrvalvalpheTyrpGlnTrpValasnglnserpheasnaIa 132
               ::::  |||  |||  ::::
Db      1360 TATGTCTTTAATCCGCCAGCAGCT-----TGGTTAAAT-----CATTTG 1398
Qy      133 llevalasntyrsersasnargserglyaspthrproilethrvalarglnleu----- 150
               :::::||||  |||  ::  |||  ::  |||
Db      1399 CTACTAAATGTTAAATAGTCTGTACGCTCTAATATTATGTACTTGGTCTTGATG 1458
Qy      151 -----glyThrAlatyrvalserAla 157
               |||  |||||  |||
Db      1459 GGAGCTATGATGCCAATGACATGGCGGCTCCATTTAATAAGACAGCTTATGTATTGCT 1518
Qy      158 ThrThrGly----- 160
               |||
Db      1519 ACTGCTGCATTAACGGAAGAAATGCAGCTCCTATTACTGCAGCGATGATGGTGTATG 1578
Qy      161 -----AlaValAlaThrAla----- 165
               |||:::|||||  |||
```

```

Db      1579 ATTCTCTCCATTAGCTATCGCTACAGCAATGTTAATCTTTAGAAAGAAATTCACCTAAAGAG 1638
Qy      166 -----leuglyleuylsSerleuthrlyshisIeu 175
               :::::||||  :::::
Db      1639 CAAAGAGGCTCAATTGTACCAAACTATGTCATGGATGATATCATTTATTACTGAAGGTGCC 1698
Qy      176 ProProleuValGlyargpheValProPheAlaAlaValAlaAlaAlaasnCysIleasn 195
               :::::|||||  |||
Db      1699 -----ATCCATTCGCAGCTGCCAT----- 1719
Qy      196 lleProleuMetArgGlnarggluleuglnvalglyileProvalAlaaspGlnAlagly 215
               |||||  ::  |||||  |||||
Db      1720 ---CCATTACGTGTGATTCCTTCATGATGTTGGCTCAGGTGTGGCTGAGCAATTGCT 1776
Qy      216 GlnArgleuGlyTyrSerValThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 233
               |||||  |||::  |||  |||||  |||
Db      1777 TTAGGTTTAAAGATCAAGTATCAAGCAACCTCATGTGTGAATTTTCGTCATTATTGCTACT 1836
Qy      234 ---Ileserargilecysmetalaleproalamealaleproleuile 250
               ::  |||  ::  |||::  |||
Db      1837 GATTTTAACCATATACTTCAACTCTTATTGCTTTAGTTGTGGTACGTTAGTT 1890

RESULT 7
US-09-252-991A-1852
; Sequence 1852, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1852
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-1852

Alignment Scores:
Pred. No.:      0.0654      Length:      669
Score:          88.00     Matches:      46
Percent Similarity: 34.96%  Conservative: 40
Best Local Similarity: 18.70%  Mismatches: 64
Query Match:    6.54%      Indels:      96
DB:             4         Gaps:      11

US-09-990-415a-2 (1-261) x US-09-252-991A-1852 (1-669)
Qy      35  ThrAspProArGAsnLeuLeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleVal 54
               :::::||||  :::::
Db      95  AGCGAGCCGCTGTGTCAGATCGTCACCGATGGCTCTACCCCAATACCGCCAACCTACGTC 154
Qy      55  GlnAsnTyrArgAlaGlyValValThrProGlyIleThrGlnAspGlnLeuTrpArgAla 74
               :::::||||  |||||  |||
Db      155  GAGAACTACGCGCGCGGTGTC-----CAGTTCTGGCGCGCC 193
Qy      75  LysTyrValTyrAspSerAla----- 81
               |||  |||
Db      194  GGG-----CTGATGTGCGGCGCGCGCGCAAGCCGTCATGCTGAGCCGCGCTACTGG 247
Qy      82  PheHisProAspThrGlyGlnLysValValleuIleGlyArgMetSerAlaGlnValPro 101
               |||:::||||  ::  |||||
Db      248  TTCACACCCGAGCTGGAAGCGCGCGGCACTGAT-----CCGCGCGCAATCGCC 298
Qy      102  MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
               ::  |||||  |||  |||||
Db      299  ATCGTCATGACCATCATCGGACCACTGCTG-----ACTGCACGTGGTGTG 343
```

```
QY      122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
      DB      344 GCGCGCAATGGAGAGCGCGCACCATGAGGACAGTGTCTCC----- 385
QY      142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
      DB      386 ---ACGCCAGCCTCGGTGGCCGAATT----- 409
QY      162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArg 181
      DB      410 -----CTGATCGGCAAG 421
QY      182 PheValProPheAlaAlaVal-----AlaAlaAlaAsnCysIle 194
      DB      422 CTGCTGCGCTACTTCTGACTCGGCATGCTGCCAGCTGGGTGGCGCGCGCTGGCGGTG 481
QY      195 AsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIleProValAlaAspGluAla 214
      DB      482 TTCGTG-----TTCGGCGCTCCGATG----- 502
QY      215 GlyGlnArgLeuGlyTyrSerValThrAlaAlaLysGlnGlyIlePheGlnValValIle 234
      DB      503 -----CGCGGCTCCTTGCTGTCCCTGTGTGCTC 529
QY      235 SerArgIleCysMetAlaIleProAlaMetAlaIleProProLeuIleMetAspThrLeu 254
      DB      530 TTGTCGCGCGGTGTTCATGTTGCGCGCGCTGGGT---CAGGTTCTGTGATTTCCTCATTTG 586
QY      255 GluLysLysAspPheLeu 260
      DB      587 GCGCGCAACAGTTCCTG 604
```

RESULT 8

```
US-09-252-991A-1931
; Sequence 1931, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1931
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1931
```

Alignment Scores:

Pred. No.:	0.169	length:	1185
Score:	88.00	Matches:	46
Percent Similarity:	34.96%	Conservative:	40
Best Local Similarity:	18.70%	Mismatches:	64
Query Match:	6.54%	Indels:	96
DB:	4	Gaps:	11

US-09-990-415A-2 (1-261) x US-09-252-991A-1931 (1-1185)

```
QY      35 ThrAspProArgAsnLeuLeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleVal 54
      DB      415 AGCGAGCGCGTGTGTCAGATGTCACCGATGGTCTCTACCCCATATACGCCCACTACGTC 474
QY      55 GlnAsnTyrArgAlaGlyValValThrProGlyIleThrGluAspGlnLeuTrpArgAla 74
      DB      475 GAGAACTACGCCCGCGCGGTGTC-----CAGTTCTGGCGCGCC 513
```

```
QY      75 LysTyrValTyrAspSerAla----- 81
      DB      514 GGG-----CTGCATGTCCGGCGCGCCGCAAGCCGTCATGCTGAGCCGCGCTACTGG 567
QY      82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101
      DB      568 TTCAAACCCGAGCTGGAAGCCGCGCGCACTGATT-----CCCGCGGCATCGCC 618
QY      102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
      DB      619 ATGTCATGACCATCATCTCGCACCATGCTG-----ACTGCACGTGTGTC 663
QY      122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
      DB      664 GCGCGCAATGGAGAGCGCGCACCATGAGGACAGTGTCTCC----- 705
QY      142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
      DB      706 ---ACGCCAGCCTCGGTGGCCGAATT----- 729
QY      162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArg 181
      DB      730 -----CTGATCGGCAAG 741
QY      182 PheValProPheAlaAlaVal-----AlaAlaAlaAsnCysIle 194
      DB      742 CTGCTGCCGCTACTCTGCTACTCGGCATGCTGCCAGCTGGGTGGCGCGCTGGCGGTG 801
QY      195 AsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIleProValAlaAspGluAla 214
      DB      802 TTCGTG-----TTCGGCGTCCGATG----- 822
QY      215 GlyGlnArgLeuGlyTyrSerValThrAlaAlaLysGlnGlyIlePheGlnValValIle 234
      DB      823 -----CGCGGCTCCTTGCTGTCCCTGTGTGCTC 849
QY      235 SerArgIleCysMetAlaIleProAlaMetAlaIleProProLeuIleMetAspThrLeu 254
      DB      850 TTGTCGCGCGGTTCATGTTGCGCGCGCTGGGT---CAGGTTCTGTGATTTCCTCATTTG 906
QY      255 GluLysLysAspPheLeu 260
      DB      907 GCGCGCAACAGTTCCTG 924
```

RESULT 9

```
US-09-252-991A-2024/c
; Sequence 2024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2024
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2024
```

Alignment Scores:

Pred. No.:	0.117	length:	552
Score:	85.00	Matches:	39
Percent Similarity:	34.76%	Conservative:	26
Best Local Similarity:	20.86%	Mismatches:	48
Query Match:	6.32%	Indels:	74
DB:	4	Gaps:	8

```
US-09-990-415A-2 (1-261) x US-09-252-991A-2024 (1-552)
QY      35 ThrAspProArgAsnLeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleVal 54
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      486 AGCGAGCCGCTGGTGCAGATCGTCACCGATGGCTCCTACCCCATATACCGCACTACGTC 427
QY      55 GlnAsnTyrArgAlaGlyValValThrProGlyIleThrGluAspGlnLeuTyrArgAla 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      426 GAGAACTACGCGCGCGGTGGTC-----CAGTTCTGGCGCGGCC 388
QY      75 LysTyrValTyrAspSerAla----- 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      387 GGG-----CTGGATGTCGCGCGCGCGCGCGCAAGCCGTCATGCTGAGCGCGCTACTG 334
QY      82 PheHisProAspThrGlyGlyValValLeuIleGlyArgMetSerAlaGlnValPro 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      333 TTCACCCCGAGCTGGAAAGCCGCGCGCACTGAT-----CCCGCGCCATCGCC 283
QY      102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      282 ATCGTCATGACCATGTCACGACCATGCTG-----ACTGCACCTGGTGGTG 238
QY      122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSergly 141
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      237 GCGCGCGAATGGAGCGCGCGCACCATGGAGGCGAGTCTGTC----- 196
QY      142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      195 ---ACGCCAGCCTCGGTGGCCGAAAT----- 172
QY      162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      171 ----- 160
QY      182 PheValProPhe-----AlaAlaValAlaAlaAla 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      159 CTGCTGCCGTACTTCGTACTCGGCATGCTGTSCAGCTGGGTGCGCGCGCTGGCGGTG 100
QY      192 AsnCysIleAsnIleProLeu 198
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      99 TTCGTGTTCGGCGTCCCGATG 79
RESULT 10
US-07-923-976-5
; Sequence 5, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikiko
; TITLE OF INVENTION: DNA Encoding Granulocyte
; TITLE OF INVENTION: Colony-Stimulating Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923, 976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
```

```
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2855 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 170..2482
; US-07-923-976-5
Alignment Scores:
Pred. No.: 2.07 Length: 2855
Score: 84.50 Matches: 54
Percent Similarity: 38.86% Conservative: 28
Best Local Similarity: 25.59% Mismatches: 76
Query Match: 6.28% Indels: 54
DB: 1 Gaps: 11
US-09-990-415A-2 (1-261) x US-07-923-976-5 (1-2855)
QY      61 ValValThrProGlyIleThrGluAspGlnLeuThrArgAlaLys-----TyrValTyr 78
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1676 ATCGTACTCCC---TTGTACCAGGACACCATGGAGCCCTGCCAGCATGTCTATGCCCTAC 1732
QY      79 AspSerAlaPheHisProAspThrGlyGlyValValLeu-----IleGlyArgMet 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1733 TCTCAAGAAATGGCTCCCTCCCATGCCCCAGAGCTGCATTAAGACATTTGGCAAGACC 1792
QY      97 SerAlaGln-----ValPrometAsnMetThrIleThrGlyCysMetLeuThrPhe 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1793 TGGGCACAGCTGGAGTGGGTGCTGAGCCCCCTGAGCTGGGGAAGAGCCCCCTTACCCAC 1852
QY      114 TyrArgLysThrProThrValValPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1853 TAC-----ACCATCTTCTGAGCAACGCTCAGAACCAGTCTTCGCGCC 1897
QY      133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThrValArg 148
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1898 ATCCTGAATGCCCTCCTCCGTGGCTTTGTCTCCATAGGCTGGAGCCGCGCAGTCTGTAT 1957
QY      149 GlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAlaLeuGlyLeu 168
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1958 CACATCCACCTCATGGCTGCCAGCCAGGCTGGGGCCACCAACAGTACAGTCTCACCCTG 2017
QY      169 LysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPheAlaAlaVal 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2018 ATGACCTTGACC-----CCAGCCCCAACAGGAGA----- 2047
QY      189 AlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlyLeuGlnValGlyIle 208
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2048 -----ATCCCTCTGGCCA- 2061
QY      209 ProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerVal---ThrAlaAlaLysGln 227
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2062 AGTGTCCAGACCCAGCTCAGACGACGCTGGGCTCTGGGTGCCACAAATCATGAGAGAG 2121
QY      228 GlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAlaIlePro 247
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2122 GATGCCTTCCAG-----CTGCCCGGCTTGGACGCCA 2154
```


QY 246 -----1leProleu1leMetaSpThleu 254
Db 17289 CAGTCGACGGCGGTGCTGCAGCTCGCTCCGACGATCATCTCGACGTTG 17339
RESULT 12
US-09-075-904-1
; Sequence 1, Application US/09075904
; Patent No. 5994137
; GENERAL INFORMATION:
; APPLICANT: Jacobs, et al.
; TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rothstein & Ebenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,904
; FILING DATE: May 11, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/247,901
; FILING DATE: May 23, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 96700/475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50341
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: L5 shuttle phasmid sequence
; HYPOTHETICAL: NO
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: L5 mycobacteriophage
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION: No. 5994137e
; AUTHORS:
; TITLE:
; JOURNAL:

; VOLUME:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-075-904-1
Alignment Scores:
Pred. No.: 512 Length: 50341
Score: 82.00 Matches: 61
Percent Similarity: 32.66% Conservative: 36
Best Local Similarity: 20.54% Mismatches: 92
Query Match: 6.10% Indels: 108
DB: 2 Gaps: 13
US-09-990-415A-2 (1-261) x US-09-075-904-1 (1-50341)
QY 44 GlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValVal--- 62
Db 16515 GGCCCGATGCTGACCGCGATCACCCCAACCTGCAGAACGTGGCCCTCGGCTGTGAAC 16574
QY 63 -----ThrProGlyIleThrGluAspGlnLeu 71
Db 16575 ATGCCGGGTGATCACCGACGTGATCACCCAGGCTCTGCTGACAGATCCAGAAC 16634
QY 72 TrpArgAlaLysTyrValTyrAspSerAlaPheHisProAspThrGlyGluValVal 91
Db 16635 ATCCTCACCAAG-----ACCGAGAGTTCTTCACG 16664
QY 92 LeuIleGlyArgMetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeu 111
Db 16665 GGCCCTCGC-----CCTGTGCTCGCTACCGGACGACGAGCCCTTCCTG 16706
QY 112 ThrPhe----- 113
Db 16707 ACGGTGTCCAACGCGCGCGGAACCTGTCGGCACGCTCTGCCCTGCAGAGATTG 16766
QY 114 -----TyrArgLysThrProThrValAlaPhe----- 122
Db 16767 ACCAAGGCTTCAACGACATGTCACACCGAGTCACGTCACAGGCGGTTCGAGGGTGC 16826
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArg----- 139
Db 16827 ATGCAAGGCTTTCGACAGCGCTGGGACGCTCTCAACCTGTTCACACCGGCTCATGAG 16886
QY 140 -----SerGlyAspThrProIleThrValArgGlnLeu 150
Db 16887 TCCGCTCTGACGGGATGGAGACAGCTCGCGGTCCTGTCGACGTTCAACAGGGTTG 16946
QY 151 GlyThrAlaTyrValSer-----AlaThrThrGlyAlaValAlaThr 164
Db 16947 GGAGATCTCTTCGTCCTCGCTGATGCCGGCTGACTTCGGTCTGTGATCGGCAAC 17006
QY 165 AlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro 184
Db 17007 GTCCTCGGG-----ACGCTGGGACACACAGCTCGCTCCCATCGTCACGGCGCTCACGCCG 17060
QY 185 ---PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
Db 17061 GCCTTCACAGCGCTGGCGAGC-----ACGCTGGACAGATGCTCACCGGAGCC 17108
QY 204 LeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThr 223
Db 17109 CTCCAAGCTCTGGGTCCGATCTGACTGACTGAGTTCGCTACGTTGATCGGACGACGCTGAAC 17168
QY 224 AlaAlaLysGlnGlyIle-----PheGln----- 231
Db 17169 ACGGCGCTGACGCTCTCCAGCGGATGCTGCCGCTCATGACAGAGCTTCACAGCATC 17228
QY 232 -----ValValIleSerArgIleCysMetAlaIleProAlaMetAla----- 245

DB 17229 TCCGACGTACTGTGACCACTGTGCCCCGACATCCCGCGCTGGCGACGCGCCCTCGGC 17288
QY 246 -----1leProProleuileMetasprhleu 254
DB 17289 CAGGTCCGAGCGCGGTGCTGACGCTCGCTCCGACGATCATCTCGACGTTG 17339
RESULT 13
US-09-426-436-1
Sequence 1, Application US/09426436
Patent No. 6225066
GENERAL INFORMATION:
APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
APPLICANT: Graham F. Hatfull
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/426,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705,557
FILING DATE:
APPLICATION NUMBER: US/08/057,531
FILING DATE:
APPLICATION NUMBER: 07/833,431
FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52297
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: phage genome sequence
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: not applicable.
ORIGINAL SOURCE:
ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
TISSUE TYPE: not applicable
CELL TYPE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
NAME/KEY:

LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993
US-09-426-436-1
Alignment Scores:
Pred. No.: 546
Score: 82.00
Percent Similarity: 32.66%
Best Local Similarity: 20.54%
Query Match: 6.10%
DB: 3
Length: 52297
Matches: 61
Conservative: 36
Mismatch: 92
Indels: 108
Gaps: 13
US-09-990-415A-2 (1-261) x US-09-426-436-1 (1-52297)
QY 44 GlyAlaGluLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValVal--- 62
DB 16412 GGGCCGAGGTGACCGCGATCACCCCAACCTGCAGACGTGGCCCTCGGCTCGTAAC 16471
QY 63 -----ThrProGlyIleThrGluAspGlnLeu 71
DB 16472 ATGGCCGGGTGATCACCGACGTGATCACCGAGCTCCTGCTGCACGACATCCAGAAC 16531
QY 72 TrpArgAlaLysTyrValTyrAspSerAlaPheHisProAspThrGlyGluValVal 91
DB 16532 ATCCTCACCAAG-----ACCGAGAGTCTCTTCACG 16561
QY 92 LeuIleGlyArgMetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeu 111
DB 16562 GGCCCTCGGC-----CCTGTCTGCTACCGGACGACGCGTTCCTG 16603
QY 112 ThrPhe----- 113
DB 16604 ACGGTGCCAACGCGCGGCGAAGCTGTCGACGCTCTGCTCCCTGCAGAGTTTC 16663
QY 114 -----TyrArgLysThrProThrValPhe----- 122
DB 16664 ACCAAGGCTTCAACGACATGTCACGTCACGTCACAGGCGGTTCGAGGTTGCC 16723
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArg----- 139
DB 16724 ATGCAAGGCTTTCGACAGCGCTGGGACGCTCTCAACCTGTTCACCGGCTCATGGAG 16783
QY 140 -----SerGlyAspThrProIleThrValArgGlnLeu 150
DB 16784 TCCGCTTCGACGCGATGGAGACCTCGCGGCTCGCTGTCACGTTTCATCAACGGGTTTC 16843
QY 151 GlyThrAlaTyrValSer-----AlaThrThrGlyAlaValAlaThr 164
DB 16844 GGAGATCTCTGCTGCTGATGCCGCGCTGACTTGGTCTGTGATCGGAAC 16903
QY 165 AlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro 184
DB 16904 GTCCCTCGGG-----ACGCTGGCACACACAGCTCGCTCCATGCTCACGCGCTACGCGC 16957
QY 185 --PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
DB 16958 GCCTTCAGACGCTGGGAGC-----ACGCTGGCACAGATGCTCACCGGAGCC 17005
QY 204 LeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThr 223
DB 17006 CTCCAAGCTCTGGGTCCGATCCTGACTCAGTCCGCTGATCGGACGACGCTGAAC 17065

QY	224	AlaAlaLysGlnGlyIle-----PheGln-----	231
Db	17066	ACGGCGCTGCAGGCTCTCCAGCCGATGCTCCGTCGCTCATGCAGAGCTTCACGACATC	17125
QY	232	-----ValValIleSerArgIleCysMetalaleProalaMetala-----	245
Db	17126	TCCGACGACTGTGACACAGTCTGGCCCCGACATCCCGGGCTGGCGACGGCCCTCGGC	17185
QY	246	-----IleProPheIleIleMetAspThrIeu	254
Db	17186	CAGTCGACGGCGGCTGCTGCAGCTTCCTCCGACGATCATCTCGACGTTG	17236
RESULT 14			
US-08-705-557-1			
; Sequence 1, Application US/08705557			
; Patent No. 6300061			
GENERAL INFORMATION:			
APPLICANT: William R. Jacobs, Jr.			
APPLICANT: Barry R. Bloom			
APPLICANT: Graham F. Hatfull			
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC			
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES			
NUMBER OF SEQUENCES: 1			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Amster, Rothstein & Eberstein			
STREET: 90 Park Avenue			
CITY: New York			
STATE: New York			
COUNTRY: U.S.A.			
ZIP: 10016			
COMPUTER READABLE FORM:			
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette			
COMPUTER: IBM PC Compatible			
OPERATING SYSTEM: MS-DOS			
SOFTWARE: Word Processor (ASCII)			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/705,557			
FILING DATE:			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US/08/057,531			
FILING DATE:			
APPLICATION NUMBER: 07/833,431			
FILING DATE: February 7, 1992			
ATTORNEY/AGENT INFORMATION:			
NAME: Pasqualini, Patricia A.			
REGISTRATION NUMBER: 34,894			
REFERENCE/DOCKET NUMBER: 96700/238			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (212) 697-5995			
TELEFAX: (212) 286-0854 or 286-0082			
TELEX: TWX 710-581-4766			
INFORMATION FOR SEQ ID NO: 1:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 52297			
TYPE: nucleotide			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE:			
DESCRIPTION: phage genome sequence			
HYPOTHETICAL: no			
ANTI-SENSE: no			
FRAGMENT TYPE: not applicable.			
ORIGINAL SOURCE:			
ORGANISM: mycobacteriophage L5			
STRAIN: not applicable			
INDIVIDUAL ISOLATE: L5			
DEVELOPMENTAL STAGE: not applicable			
HAPLOTYPE: not applicable			
TISSUE TYPE: not applicable			
CELL TYPE: not applicable			
CELL LINE: not applicable			
ORGANELLE: not applicable			

```

1 IMMEDIATE SOURCE: mycobacteriophage L5 particles
2 POSITION IN GENOME: entire genome
3 FEATURE:
4 NAME/KEY:
5 LOCATION:
6 IDENTIFICATION METHOD:
7 OTHER INFORMATION:
8 PUBLICATION INFORMATION:
9 AUTHORS: Hatfull and Sarkis
10 TITLE: DNA Sequence, Structure and Gene
11 TITLE: Expression of Mycobacteriophage L5:
12 TITLE: A Phage System for Mycobacterial
13 TITLE: Genetics
14 JOURNAL: Molecular Microbiology
15 VOLUME: 7
16 PAGES: 395-405
17 DATE: 1993
18
19 US-08-705-557-1

```

Alignment Scores:		
Pred. No.:	546	Length: 52297
Score:	82.00	Matches: 61
Percent Similarity:	32.66%	Conservative: 36
Best Local Similarity:	20.54%	Mismatches: 92
Query Match:	6.10%	Indels: 108
DB:	4	Gaps: 13

US-09-990-415A-2 (1-261) x US-08-705-557-1 (1-52297)

44 GLYALGLNLEUGLIALSERARGASNILEVALGLNASNTYRRARGALGLYVALVAL --- 62

Db 16412 GGCCCGATGCTGACCGCGATCACCCCAACCTGCAGAACGTGGCCTCGGGCCTGTGAAC 16471

QY 63 -----ThrProGlyIleThrGluAspGlnLeu 71

Db 16472 ATGGCCGGTGCATCACCGACGTGATCACCCAGGCTCTGGTCTGCAGCAGATCCAGAAC 16531

72 tTpaTgaLaLystYvaLTytsrSeraLaPhenIsProasrThGlgUlsvalval 91

Db 16532 ATCCTCACCAG-----ACCGAGAGTTCTTCACG 165

92 LeuIleGlyargmetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeu 111

Db 16562 GGCCTCGGC-----CCTGTGCTCGCTACCGGGCAGCGGCTTCCTG 16603

qy	112	Thrphe	113
----	-----	--------	-----

Db 16604 ACGCTGTCCAACGCCGGCGGAATCGTTCGGCACGCTCCTGGCTCCCTGCAGGAGTTC 16663

QY 114 -----TyrArgLysThrProThrValValAlphe----- 122

Db 16664 ACCAACGGCTTCAACGACATGGTCAACCGAGTCACGTCCAACGGCGGTGTTCCGAGGGTCCC 16723

123 TPGINTpvalasnglInserpheasnaIaIlevalasntYrSerasnarg----- 139

Db 16724 ATGCAAGGGCTTTCGCAGACGCTGGGCAGCGTCCCTCAACCTGTCAACCGGCTCATGGAG 16783

QY 140 -----SercGlyaspThrproIethrValArgcInleu 150
 ||| : : :

Db 16784 TCCGGTCTGCAGCGGATGGACAGCTCGGGGTTCCGCTGTGCAGCTTCATCAACGGGTTc 16843

QY 151 GlyThrAlatyrValSer-----AlaIthrThrgIyAlaValAlaIthr 164
 ||| :||| ||| :||| :||| ||| :|||

DB 16844 GGAGATCTCTTCGTCCTCGCTGATGCCGGCCGCTGACTTCGGTCTCTGGTCTGATCGGCAAC 16903

105 A L A L E U G I L L E U L y s s e r L e u I n L y s N I S L e u P r o P r o L e u v a l G I y A t r p h e n a l P r o 184

DB 18904 GTCCTCGGG-----ACGCTGGGCACACAGCTCCCTCCCACTGTCACGGCGCTCACGGCG 18957

QY 185 ---PHEALDAVAIAIAIAIASNCYSIIeashIlePTOLEumetaRgInatrgId 203

||| ::|||::: |||

DB 18958 GCGTTCAGACGCTGGGAGC-----ACGCTCGGACGATGCTCACCAGAGCC 17005

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 14:00:26 : Search time 263 Seconds
(without alignments)
2678.910 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGEPLPDINIQPRWD.....PAMAIPPLIMDTLEKKDFLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09990415/runat_16102003_073806_9430/app_query.fasta_1.455
-DB=N_Geneseq_19Jun03 -QMT=fastlap -SUFFIX=rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415@cgn2_1_1-396_runat_16102003_073806_9430 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1345	100.0	1232	24	ABN83754	Human protein clus
2	1345	100.0	1375	24	ABO98865	Human ORE672 codin
3	1345	100.0	1729	22	AAI59123	Human polynucleoti
4	1345	100.0	1826	22	AAF27703	Human transport pr
5	1345	100.0	2702	22	AAK94713	Human full-length
6	1345	100.0	3047	22	AAH72722	Human cervical can
7	1318	98.0	984	24	ABL39805	Human NS CDNA sequ
8	1269	94.3	843	21	AAC75798	Human OREFX ORF1353
9	1223.5	91.0	2215	23	AAS84362	DNA encoding novel
10	1089	81.0	1076	22	AAI59569	Human polynucleoti
11	1089	81.0	1134	24	ABN84365	Alzheimer's diseas
12	1089	81.0	1485	22	AAL26555	Human breast cance
13	1089	81.0	2053	22	AAI61354	Human polynucleoti
14	1089	81.0	2269	24	ABN83757	Human protein clus
15	1089	81.0	2367	24	AAD21995	Human transporters
16	1089	81.0	2468	21	AAC75794	Human OREFX ORF1349
17	1082	80.4	2947	22	AAH44832	Rat CDNA encoding
18	1053	78.3	752	22	AAK91763	Human CDNA 5'-end
19	1053	78.3	752	22	AAK93199	Human CDNA clone r
20	1051.5	78.2	1503	24	ABK54156	CDNA encoding huma
21	1026	76.3	2766	22	AAH15610	Human CDNA sequenc
22	1003.5	74.6	1061	24	ABN83755	Human protein clus
23	961.5	71.5	1084	24	ABK96715	DNA encoding splic
24	897.5	66.7	984	24	ABK96716	Human encoding Alz
25	897.5	66.7	984	24	ABN84366	Alzheimer's diseas
26	798	59.3	2642	22	AAH17727	Human CDNA sequenc
27	765	56.9	741	22	AAH06089	Human CDNA clone (
28	754	56.1	1286	23	ABL18135	Drosophila melanog
29	754	56.1	1567	24	ABN83756	Human protein clus
30	750	55.8	1772	24	AAD33666	Human TRICH-21 CDN
31	750	55.8	2312	22	AAI59647	Human polynucleoti
32	750	55.8	2376	22	AAI57861	Human polynucleoti
33	750	55.8	2458	24	ABK35352	Human CDNA encodin
34	712.5	53.0	1575	24	ABL99804	Human secretory po
35	693	51.5	806	22	AAH04784	Human CDNA clone (
36	670	49.8	507	23	ABV45979	Human prostate exp
37	626	46.5	421	23	ABV16181	Human diagnostic a
38	602	44.8	524	22	AAS31167	Drosophila melanog
39	599.5	44.6	1585	23	ABL11787	Drosophila melanog
40	578	43.0	4258	23	ABL18134	Drosophila melanog
41	563	41.9	4438	23	ABL11786	Drosophila melanog
42	545	40.5	413	23	AAS84361	DNA encoding novel
43	480.5	35.7	2673	24	AAD36304	Human transporter
44	467	34.7	1244	20	AAK97798	Extended human sec
45	462	34.3	1548	24	ABL90501	Human polynucleoti

ALIGNMENTS

RESULT 1
ABN83754
ID ABN83754 standard; CDNA: 1232 BP.
XX
AC ABN83754;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human protein cluster I gene.
XX
KW Protein cluster I; human; metabolic disorder; obesity; diabetes;
XX antidiabetic; diagnosis; therapy; chromosome 10; gene; ss.
OS Homo sapiens.
XX
FH Key
FT CDS 450..1232 Location/Qualifiers

```
FT      /tag= a
FT      /product= "Protein cluster I"
FT      /partial
FT      /note= "the CDS does not include a stop codon"
PN      WO200242324-A1.
XX
PD      30-MAY-2002.
XX
PF      22-NOV-2001; 2001WO-SE02581.
XX
PR      24-NOV-2000; 2000SE-0004325.
XX
PA      (PHAA ) PHARMACIA AB.
XX
PI      Altersand A;
XX
DR      WPI; 2002-500277/53.
DR      P-PSDB; ABB76445.
XX
PT      Novel nucleic acid molecule encoding Protein Cluster I, useful in the
PT      diagnosis of metabolic diseases, such as obesity and diabetes, and in
PT      the identification of agents useful in the treatment of the diseases
XX
PS      Claim 1; Page 23-24; 34pp; English.
XX
XX      The present invention relates to the identification of a human gene
CC      family (see ABN83754-57) encoding a group of polypeptides (see
CC      ABB76445-48) referred to as Protein Cluster I. This family of
CC      homologous proteins was identified by an 'all-versus-all' BLAST
CC      procedure using all Caenorhabditis elegans proteins in the
CC      Wormpep20 database release. Proteins of unknown function were
CC      compared to the Drosophila melanogaster Flybase database, and
CC      non-annotated protein clusters, conserved in both C. elegans and D.
CC      melanogaster, were used in a BLAST procedure against the Celera
CC      Human Genome Database, and Protein Cluster I proteins of unknown
CC      function were selected for study. The human part of Protein
CC      Cluster I comprises polypeptides encoded by 3 genes; an alternative
CC      splicing of the present sequence (deletion of nucleotides 624-794)
CC      gives the sequence in ABN83755. The present gene was identified in
CC      a human DNA sequence from clone RP11-108L7 on chromosome 10. It
CC      is expressed primarily in the nervous system and digestive system.
CC      The claimed Protein Cluster I nucleic acid molecules and proteins
CC      are proposed to be useful for differential identification of the
CC      tissue(s) or cell type(s) present in a biological sample, for
CC      diagnosis of diseases and disorders, including metabolic disorders
CC      and immune disorders, especially obesity and diabetes, and for
CC      identifying agents useful in the treatment of such diseases. The
CC      nucleic acid molecules are also useful as hybridisation probes, for
CC      chromosome and gene mapping, in PCR technologies, and in the
CC      production of sense or antisense nucleic acids.
XX
SQ      Sequence 1232 BP; 239 A; 368 C; 375 G; 250 T; 0 other;
Alignment Scores:
Pred. No.:      8.01e-143      Length:      1232
Score:      1345.00      Matches:      261
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      24      Gaps:      0
US-09-990-415a-2 (1-261) x ABN83754 (1-1232)
QY      1 MetGlusErlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTTPasp 20
DB      450 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 509
QY      21 GlnSerThrphelGlyValArgAlaArgHisphetheValThrAspProArgAsnLeu 40
DB      510 CAAAGTACTTCTCTGGGAGAGCCCGGACATTCTTCACTGTACTGATCCCTGAAATCTG 569
QY      41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
```

```
DB      570 CTGCTGTCCGGGGGCACAGCTTGAAGCTTCTCGGAACATCGTCAGAACTACAGGGCCGGC 629
QY      61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
DB      630 GTGGTGACCCCAAGGATCACCGAGACAGCAGCTGTGAGAGGCCAAGTATGTATGACTGCC 689
QY      81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB      690 GCCTTCCATCCGGACACAGGGAGAGAGGTGTCTCTGATTTGGCCGATGTACAGCCAGTTC 749
QY      101 PrometasmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB      750 CCCATGAACATGACCATCACATGCGTGCATGCTCACATTCACAGGAAGACCCCAACCTGC 809
QY      121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB      810 GTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATTGTTAAGTACTCCAACCCGAGT 869
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB      870 GGTAGACACTCCCATCACTGTGAGGACAGCTGGGGACAGCCTATGTGAGTGCCACCACTGGA 929
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB      930 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACTGCCCTTGCTGGC 989
QY      181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB      990 AGATTGTGCCCTTTGACAGCAGTGAGTGACGTGCCAACTGCATCAACATCCCTGATGAGG 1049
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB      1050 CAGAGAGAGCTGACAGGTGGGCATCCCGGTGGCTGATGAGGAGGTGAGAGGCTTGCTTAC 1109
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB      1110 TCGGTGACTGCAGCCAAAGCAGGAAATCTTCCAGGTGTGTGAATTCAAGAATCTGCATGCG 1169
QY      241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
DB      1170 ATTCCCTGCCATGGCCATCCCACTGATCATGACACACTGTGAGAAGAAGACTTCTTG 1229
QY      261 Lys 261
DB      1230 AAG 1232
RESULT 2
ABQ98865
ID      ABQ98865 standard; DNA; 1375 BP.
XX
AC      ABQ98865;
XX
DT      04-NOV-2002 (first entry)
XX
DE      Human ORF672 coding sequence.
XX
KW      Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW      Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW      human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW      cancer; cardiovascular disease; allergy; autoimmune disease;
KW      wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS      Homo sapiens.
XX
PN      US2002082206-A1.
XX
PD      27-JUN-2002.
XX
PF      30-MAY-2001; 2001US-0867550.
XX
PR      30-MAY-2000; 2000US-208427P.
```

PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI: 2002-626554/67.
DR P-PSDB; ABP64302.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease
XX
XX Claim 2; SEQ ID 1343; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer, blood
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/sequence.html?docid=20020082206.
XX
SQ Sequence 1375 BP; 342 A; 371 C; 374 G; 286 T; 2 other;

Alignment Scores:
Pred. No.: 9.4e-143 Length: 1375
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-990-415A-2 (1-261) x ABQ98865 (1-1375)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIuProArgTrpAsp 20
Db 91 ATGGAAGCAAAATGGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGTGGGAC 150

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 151 CAAAGTACTTCTCTGGGACAGAGCCCGGCACTTTTCACTGTACTGATCTCGAATCTG 210

QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 211 CTGCTGTCCGGGGACACAGCTGGAAGCTTCTCGGAACATCGTGCAAGTACAGGCGGC 270

QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaIleTyrValTyrAspSer 80
Db 271 GTGGTGACCCAGGAGATCACCAGGACCACTGTGAGAGGCCCAAGTATGTATGACTCC 330

QY 81 AlaPheHisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 331 GCCCTTCATCCGGACACACAGGGGAGAGGTGGTCTGATTTGGCCGCATGTACGCCAGGTG 390

QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 391 CCCATGAACATGACCATCATCTGCTGCATGCTCACAATTCTACAGAGAACCCCAACCGTG 450

QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 451 GTGTTTGGCAGTGGGTGAATCAGTCTCTCAATGCATTTGTTAACTACTCCCAACCGCAGT 510

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160

Db 511 GGTGACACTCCCACTACTGTGAGGCAGCTGGGGACAGCCTATGTGAGTGCCACCACTGGA 570
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 571 GCTGTGGCAGCGGCGCTGGGACTCAATCCCTCACCAAGCACCTGCCCCCTTGTCGGC 630
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200
Db 631 AGATTGTGCCCTTTGCAGCAGCATGTGGCAGCTGCCAAGCTGCATCAACATCCCTCATGAGG 690
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 691 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAGGTCAGAGGCTTGCTTAC 750
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 751 TCGGTACTGCAGCCAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGAATCTGCATGGCG 810
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 811 ATTCTGTCATGGCCATCCCACTGATCATGACACTCTGAGAGAAGAAGACTTCTTG 870
QY 261 Lys 261
Db 871 AAG 873

RESULT 3
AAI59123
ID AAI59123 standard; cDNA; 1729 BP.
XX
AC AAI59123;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1326.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB; AAM39967.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

PS Claim 1; SEQ ID NO 1326; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence: 1729 BP; 406 A; 469 C; 491 G; 363 T; 0 other;

Alignment Scores:

Pred. No.:	1.31e-142	Length:	1729
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-990-415A-2 (1-261) x AA159123 (1-1729)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpasp 20
 Db 247 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGAACCTCGCTGGGAC 306
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 Db 307 CAAAGTACTTCTCGGGCAGAGCCCGGCACCTTTTCACTGTACTGATCCTCGAATCTG 366
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 Db 367 CTGCTGTCCGGGGCAGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCGGC 426
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80
 Db 427 GTGGTGACCCCAAGGATCACCGAGGACAGCTGTGGAGGGCCAAATATGTATGACTCC 486
 QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 Db 487 GCCCTTCATCCGACACAGGGGAGAAAGGTGCTCCTGATTTGGCCGATGTACAGCCCGGTG 546
 QY 101 PrometasMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 Db 547 CCCATGAACATGACCATCTGCTGCTGCTACATCTTACAGAGAAGACCCCAACCGGTG 606
 QY 121 ValPheTrpGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 Db 607 GTGTTCAGAGTGGGTGATCATCTCCTCATATGCCATTTGTTACTACTCCAAACCGCAGT 666
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 Db 667 GGTGACACTCCCATCTGAGTGAAGGAGCTGGGGACAGCCTATGTGAGTGCCACCACTGGA 726
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 Db 727 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACTGCCCCCTTGGTGGCC 786
 QY 181 ArgPheValProPheAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 Db 787 AGATTGTGCTTTCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGTATGAGG 846
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220

Db 847 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGGAGCTCAGAGGCTTGCTAC 906
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
 Db 907 TCGGTGACTGCAGCCCAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGATCTGCATGGCG 966
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 Db 967 ATTCTGCCATGGCCATCCCAACCACTGATCATGTGACACTCTGGAGAAGAAGACTTCTG 1026
 QY 261 Lys 261
 Db 1027 AAG 1029

RESULT 4

AAAF27703
 ID AAF27703 standard; cDNA; 1826 BP.

XX AAF27703;

XX AC AAF27703;

XX DT 28-MAR-2001 (first entry)

XX DE Human transport protein TPPT-3 coding sequence.

XX KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer; ss.

XX OS Homo sapiens.

XX PN WO200078953-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US16668.

XX PR 17-JUN-1999; 99US-0139923.

XX PR 10-AUG-1999; 99US-0148177.

XX PR 18-AUG-1999; 99US-0149357.

XX PR 28-OCT-1999; 99US-0162287.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;

XX PI Baughn MR, Azimzai Y, Lu DM, Au-Young J, Patterson C;

XX DR WPI; 2001-041424/05.

XX DR P-PSDB; AAB60083.

XX PT Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -

XX PS Claim 5; Page 142-143; 165pp; English.

XX CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.

XX SQ Sequence 1826 BP; 415 A; 519 C; 512 G; 380 T; 0 other;

Alignment Scores:

Pred. No.:	1.42e-142	Length:	1826
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-990-415A-2 (1-261) x AAF27703 (1-1826)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 449 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 508
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 509 CAAAGTACTTCTCTGGGACAGCCGGCACTTTTTCACCTGTACTGATCCCGAAATCTG 568
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 569 CTGCTGTCCGGGGCAGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGGC 628
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrParGAlaLysTyrValTyrAspSer 80
Db 629 GTGGTGACCCAGGAGGATCACCAGAGCAGCTGTGAGGGCCCAAGTATGTATGACTCC 688
QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 689 GCCTTCATCCGGACACAGGGGAGAGGTGGTCTGATTGGCCGATGTACGCCACAGGTG 748
QY 101 PrometasMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 749 CCCATGAACATGACCATCACTGCTGCTGATCTACAGGAAGACCCCAACCCGTG 808
QY 121 ValPheTyrGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 809 GTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCGCATGTGTAACCTCAACCCGAGT 868
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 869 GGTGACACTCCCATCACTGAGGAGCTGGGGACAGCCCTATGTAGTGCACCACTGGA 928
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 929 GCTGTGGCCACGGCCCTGGAGCTCAAAATCCCTCACCAAGCACCTGCCCTGTGTCGGC 988
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 989 AGATTGTGCCCTTGTGACGACGTGGCAGCTGCCAATGCATCAATCCCCCTGATGAGG 1048
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 1049 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGACGTGAGAGGCTTGCTG 1108
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 1109 TCGGTGACTGCAGCCAAAGCAGGGAATCTTCCAGGTGTGATTCAAGAACTGATGGCG 1168
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
Db 1169 ATTCCTGCATGCGCATCCACCACTGATCATGACACTCTGGAGAAGAAAGACTTCTCTG 1228
QY 261 Lys 261
Db 1229 AAG 1231

RESULT 5
AAK94713
ID AAK94713 standard; cDNA; 2702 BP.
XX AAK94713;
AC
XX 06-NOV-2001 (first entry)
DT
XX Human full-length cDNA, SEQ ID NO: 3751.
DE
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW
XX Homo sapiens.
OS
XX EP1130094-A2.
PN
XX 05-SEP-2001.
PD

XX 07-JUL-2000; 2000EP-0114089.
PF
XX 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR MPI; 2001-524255/58.
DR P-PSDB; AAM93760.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3751; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 2702 BP; 608 A; 757 C; 699 G; 638 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2,51e-142 Length: 2702
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
US-09-990-415A-2 (1-261) x AAK94713 (1-2702)
QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 70 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCTGGGAC 129
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 130 CAAAGTACTTCTCTGGGACAGAGCCGGCACTTTTTCACCTGTACTGATCTCGAAATCTG 189
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 190 CTGCTGTCCGGGGCAGACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGGC 249
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrParGAlaLysTyrValTyrAspSer 80
Db 250 GTGGTGACCCAGGAGGATCACCAGAGCAGCTGTGAGGGCCCAAGTATGTATGACTCC 309
QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 310 GCCTTCATCCGGACACAGGGGAGAGGTGCTCTGATGGCCGCACTGACAGCCAGGTG 369
QY 101 PrometasMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 370 CCCATGAACATGACCATCACTGCTGCTGCTCATATTTCTACAGGAAGACCCCAACCGTG 429
QY 121 ValPheTyrGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 430 GTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCGCATGTGTAACCTCAACCCGAGT 489

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 490 GSTGACACTCCCATCTGACTGTGAGGCACAGCTGGGACAGCCTATGTAGTGCCACCACTGGA 549
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 550 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACTTGCCCCCTTGTCGGC 609
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 610 AGATTGTGCCCTTTGGACAGATGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAGG 669
QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyr 220
DB 670 CAGAGAGAGCTGCAGGTGGGCATCCCATGTGCTGATGAGGCAGGTCAAGAGGCTTGCTAC 729
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
DB 730 TCGGTGACTGCAGCCACAGCAGGAATCTCCAGGTGTGATTCCAAGAAATCTGCATGGCG 789
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
DB 790 ATTCTGCGCATGGCCATCCACCACTGATCATGACACTCTGGAGAAAGAACTTCCTG 849
QY 261 Lys 261
DB 850 AAG 852

RESULT 6
AAH72722/c
ID AAH72722 standard; cDNA; 3047 BP.

AAH72722;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 3996.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

WO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000WO-US33312.

08-DEC-1999; 99US-0169681.

21-DEC-1999; 99US-0171350.

14-MAR-2000; 2000US-0189315.

12-MAY-2000; 2000US-0203791.

09-JUN-2000; 2000US-0210600.

21-JUL-2000; 2000US-0220114.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

WPI; 2001-375006/39.

New isolated nucleic acid for diagnosing and treating cervical cancer

and for assessing and detecting compounds for treating the cancer -

Claim 1; Page 815-816; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with

cervical cancer with cytostatic activity. The nucleic acids and encoded

polypeptides are useful: to assess if a patient is afflicted with

cervical cancer or has a pre-malignant condition; to monitor the

progression of cervical cancer or a premalignant condition in a patient;

and to select and/or assess the efficacy of a compound or therapy for

inhibiting cervical cancer in a patient. The nucleic acids may also be

CC useful for gene therapy.
XX
SQ Sequence 3047 BP; 629 A; 790 C; 891 G; 733 T; 4 other;

Alignment Scores:
Pred. No.: 2,99e-142 Length: 3047
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-990-415a-2 (1-261) x AAH72722 (1-3047)

QY 1 MetGluSerLysMetGlyGlnLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20

DB 2605 ATGGAAGCAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAACCTCGCTGGAC 2546

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40

DB 2545 CAAAGTACTTTCCTGGGACAGACCCGGCACTTTTCACGTGTACTGATCCTCGAATCTG 2486

QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgIleGly 60

DB 2485 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGCGGGC 2426

QY 61 ValValThrProGlyIleThrGlnAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80

DB 2425 GTGGTGACCCCAAGGATCAACCGAGACCAAGCTGTGGAGGGCCAAATATGTGTAGACTCC 2366

QY 81 AlaPheHisProAspThrGlyGlnLysValLeuIleGlyArgMetSerAlaGlnVal 100

DB 2365 GCCTTCATCCGGACACAGGGGAGAAGGTGGTCTGATTTGGCCGATGTCCAGCCAGGTG 2306

QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120

DB 2305 CCCATGAACATGACCACTCACTGGCTGCATGCTCACATTCCTACAGGAAGACCCCAACCTG 2246

QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140

DB 2245 GTGTTCTGGCAGTGGGTGAATCAGTCCCTTCAATGCCATGTGTTAACTACTCCAAACCGCAGT 2186

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160

DB 2185 GGTGACACTCCCATCTGACTGTGAGGCAGCTGGGACAGCCTATGTGAGTGCCACCACTGGA 2126

QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180

DB 2125 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCACCAAGCACACTGCCCTTGTCGGC 2066

QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200

DB 2065 AGATTGTGCCCTTTGGACAGACTGGCAGCTGCCAACTGCATCAACATCCCCCTGATGAGG 2006

QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyr 220

DB 2005 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGGCAGGTCAAGGCTTGCTAC 1946

QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240

DB 1945 TCGGTGACTGCAGCCAAGCAGGAATCTTCACGTGTGATTTCAAGAAATCTGCATGGCG 1886

QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260

DB 1885 ATTCTGCGCATGGCCATCCACCACTGATCATGACACTCTGGAAGAAGAACTTCCTG 1826

QY 261 Lys 261

DB 1825 AAG 1823

RESULT 7
ABL39805
ID ABL39805 standard; cDNA; 984 BP.

XX ABL39805;
AC
XX 10-MAY-2002 (first entry)
DT
XX Human NS CDNA sequence SEQ ID NO:115.
DE
XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200206315-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001WO-IL00653.
PF
XX 18-JUL-2000; 2000IL-0137345.
PR
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
XX Mintz L, Freilich S, Bernstein J;
PI
XX
DR WPI; 2002-155037/20.
DR
XX P-PSDB; ABB06151.
PT
XX One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 1; Page 141-142; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquiliser, antiulcer,
CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, diabetes, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SO Sequence 984 BP; 224 A; 278 C; 285 G; 196 T; 1 other;

Alignment Scores:
Pred. No.: 6, 77e-140 Length: 984
Score: 1318.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62% Mismatches: 1
Query Match: 97.99% Indels: 1
DB: 24 Gaps: 0
US-09-990-415A-2 (1-261) x ABL39805 (1-984)
QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
DB 49 ATGGAAGCAAAATGGGTGAATTGCCCTTAGACATCAACATCCAGGAACCTCGCGGAC 108
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 109 CAAAGTACTTCCCTGGGCAGAGC-CGGCACTTTTCACGTGTACTGATCCCGAAATCTG 167
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 168 CTGCTGTCGGGGCAGACGCTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGGCCGGC 227
QY 61 ValValThrProGlyIleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSer 80
DB 228 GTGCTGACCCCGAGGATCACCGAGACCACTGTGGAGGGCCAAAGTATGTATGACTCC 287
QY 81 AlaPheHisProAspThrGlyGluLysValIleLeuIleGlyArgMetSerAlaGlnVal 100
DB 288 GCCTTCATCCGACACAGGGAGAGAGGTGCTGATTGGCCGCATGTCAAGCCAGGTG 347
QY 101 PrometasmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 348 CCCATGAACATGACCATCACTGGCTGCATGCTCACAATCTACAGGAAGACCCCAACCGTG 407
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 408 GTGTTCTGGCAGTGGGTGAATCAGTCCCTCAATGCCATGTTAATACTCAACCGCAGT 467
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 468 GGTGACACTCCCATCACTGTGAGGACAGCTGGGACAGCCATGTGAGTGCACCACTGGA 527
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 528 GCTGTGGCCACGGCCCTGGGACTCAAAATCCCTCAACCAAGCACCTGCCCTTGGTCGGC 587
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 588 AGATTGTGCCCTTGCAGCAGTGGCAGCTGCCAATGCATCAACATCCCTGTGATGAGG 647
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 648 CAGAGAGAGCTGAGGTGGCATCCCGGTGGTGAAGGACAGTCAAGGCTTGGCTTAC 707
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
DB 708 TCGGTGACTGCAGCCAAAGGAGGAAATCTTCCAGGTGGTGAATTCAGAAATCTGCATGGCG 767
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
DB 768 ATTCTGCTCCATGGCCATCCCAACCACTGATGAGACACTCTGGAGAGAAGAAACTTCTG 827
QY 261 Lys 261
DB 828 AAG 830
RESULT 8
AAC75798
ID AAC75798 standard; cDNA; 843 BP.
XX
XX AAC75798;
AC
XX 08-FEB-2001 (first entry)
DT
XX Human ORFX ORF1353 polynucleotide sequence SEQ ID NO:2705.
DE
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

KW vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antihemolytic; antithyroid;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB41589.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 1942; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

XX which represent the human ORFX open reading frames 1 to 3161. The ORFX

XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;

XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;

XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;

XX antidiabetic; hypotensive; dermatological; immunosuppressive;

XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;

XX antithyroid; and antianaemic. The sequences can be used for determining

XX the presence of or predisposition to, or preventing or treating

XX pathological conditions associated with an ORFX-associated disorder. The

XX nucleic acids can be used to express ORFX proteins in gene therapy

XX vectors. The proteins and nucleic acids may be used to treat cancers,

XX proliferative disorders, neurodegenerative disorders, osteoarthritis,

XX graft vs host disease, cardiovascular disease, diabetes mellitus,

XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus

XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,

XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,

XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,

XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance

XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 843 BP; 183 A; 241 C; 252 G; 165 T; 2 other;

XX Alignment Scores:

XX Pred. No.: 2, 01e-134

XX Score: 1269.00

XX Percent Similarity: 99.20%

XX Best Local Similarity: 98.01%

XX Query Match: 94.35%

XX DB: 21

XX Gaps: 0

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
Db 91 ATGGAAGCAAAATGGGTGAATTCCTTAGACATCAACATCCAGAACCTCGCTGGGAC 150
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 151 CAAAGTACTTTCCTGGGACAGAGCCCGCAGCTTTTTCACAGTGTACTGATCCCGAAATCTG 210
QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrrArgAlaGly 60
Db 211 CTGCTGTCCGGGACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTACAGGCGCGGC 270
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrrValTyrrAspSer 80
Db 271 GTGGTGACCCCGAGGATCACCAGAGACCAAGCTGTGAGAGGCGCAAGTATGTATGACTCC 330
QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 331 GCCTTCCATCCCGACACAGGGAGAGAGGTGGTCTGATTTGGCCGATGTCAGCCAGGTG 390
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrrArgLysThrProThrVal 120
Db 391 CCCATGAACATGACCATCATCTGCTGCATGCTCACAATTCTACAGGAAGACCCCAACCGTG 450
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrrSerAsnArgSer 140
Db 451 GTGTTCTGGCAGTGGGTGAATCAAGTCTTCAATGCCAATGTTAACTACTCCAAACCGCAGT 510
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrrValSerAlaThrThrGly 160
Db 511 GGTGACACTCCCATCATCTGAGGACGAGCTGGGACAGCCTATGTGAGTGCCACCACTGGA 570
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 571 GCTGTGGCCACGCGCCGTGGAGCTCAATCCCTCACCAAGCACCTGCCCTTGGTCGGC 630
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnLysIleAsnIleProLeuMetArg 200
Db 631 AGATTGCTACCTTTGACAGCAGTGAGCAGCTGCCAATGCATCAACATCCCGCTGATGAGG 690
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 691 CAGAGGAGCTGACAGTGGGATCCCAAGTGAAGTGAAGCTGTCAGAGACTTGGCCAC 750
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
Db 751 TCGGTGACTGCTGCCAAGAGGCGCATCTTCCAGGTGTGTATCGAAGATCGGCATGGCG 810
QY 241 IleProAlaMetAlaIleProProLeuIleMet 251
Db 811 ATCCCCGCGCATGGCCATTTCCCGGTGATCATG 843

RESULT 9

AAS84362 standard; cDNA; 2215 BP.

AAS84362;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #20166.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

US-09-990-415A-2 (1-261) x AAC75798 (1-843)

XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG20175.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID NO 20166; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2215 BP; 491 A; 649 C; 582 G; 492 T; 1 other;

Alignment Scores:
Pred. No.: 1,21e-128 Length: 2215
Score: 1223.50 Matches: 256
Percent Similarity: 89.82% Conservative: 0
Best Local Similarity: 89.82% Mismatches: 2
Query Match: 90.97% Indels: 27
DB: 23 Gaps: 1

US-09-990-415A-2 (1-261) x AAS84362 (1-2215)

QY 4 lysMetgIyGluLeuProLeuAspIleAsnIleGlnIuProArgTTPaspGlnSerThr 23
DB 208 AAAATGGGTGAATGCTTTAGACATCAACATCCAGAACCTCGCTGGAGCAAAAGTACT 267
QY 24 PheLeuGIyArgAla-ArgHisPhePheThrValThrAspProArgAsnLeuLeuLeu 43
DB 268 TTCCTGGGCAGAGCGCCGGCACCTTTTCTACTGTACTGATCTCGAAATCTGCTGCTGTC 327
QY 43 rGIyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValTh 63
DB 328 CGGGGCACAGCTGGAACTTCTCGGAACATCGTGACAACTACAGGGCGCGTGTGAC 387
QY 63 rProGlyIleThrGluAspGlnLeuTrpArgAlaIlystYrValTyrAspSerAlaPheH 83
DB 388 CCCAGGATCACCGAGACAGCTGTGGAGGGCCAAAGTATGTATGACTCCGCTTCCA 447
QY 83 sProAspThrGlyGluIyValValLeuIleGlyArgMetSerAlaGlnValPrometas 103
DB 448 TCCGGACACAGGGGAGAAGGTGTCTCTGATTTGGCCCATGTCAAGCCAGGTGCCCATGAA 507
QY 103 nMetThrIleThrGlyCysMetLeuThrPheTyr----- 114

DB 508 CATGACCATCAGCTGGCTGCATGCTCACATTCATACAGGCGAGGGTCTAAAGATGAGGGCCA 567
QY 115 -----ArgLysThrProThrValValPh 122
DB 568 CTGTAGACGGGGCAGAAGTAGTGTCTTGTTCCTCAGGAAGACCCCAACCGTGTGTT 627
QY 122 eTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArg-SerGlyA 142
DB 628 CTGGCAGTGGGTGAATCAGTCTTCAATGCCATTGTTAATCTCCAACCGCAGGTGTG 687
QY 142 sPThr-ProIleThr-ValArgGlnLeuGlyThr-AlaTyrValSerAla-ThrThrGly 160
DB 688 ACACCTCCCATCACTTGTGAGGACAGCTGGGGACAGCAGCTTATGTAGTGCANCCACTGA 747
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 748 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCACCTGCCCTTGCTCGGC 807
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 808 AGATTGTGCCCCCTTGCAGACAGTGGCAGCTGCCAATGCATCAACATCCCCCTGATGAGG 867
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 868 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGCTGATGAGGCAGGTCAAAGGCTTGGCTAC 927
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
DB 928 TCGGTGACTGCAGCCAAAGCAGGGAATCTTCCAGGTGGTGTATTCTAGAAATCTGCATGGCG 987
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
DB 988 ATTCTGTCATGGCCATCCACACCATGATCATGACACTCTGAGAGAAGAAAGCTTCCTG 1047
QY 261 Lys 261
DB 1048 AAG 1050
RESULT 10
AAI59569
ID AAI59569 standard; cDNA; 1076 BP.
XX
AC AAI59569;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1772.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR MPI: 2001-442253/47.
 DR P-PSDB: AAM40413.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 1772; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1076 BP; 285 A; 258 C; 269 G; 264 T; 0 other;
 Alignment Scores:
 Pred. No.: 8.2e-114 Length: 1076
 Score: 1089.00 Matches: 201
 Percent Similarity: 89.45% Conservative: 28
 Best Local Similarity: 78.52% Mismatches: 27
 Query Match: 80.97% Indels: 0
 DB: 22 Gaps: 0
 US-09-990-415A-2 (1-261) x AA159569 (1-1076)
 QY 6 GlyGluLeuProLeuAspIleAsnIleGlnIuProArgTTPaspGlnSerThrPheLeu 25
 DB 67 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAGCACTTTCAATT 126
 QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
 DB 127 GGACGAGCCAAATCATTTCTTCACTGTACTGACCCAGAACATCTGTTAAACCAAGAA 186
 QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
 DB 187 CAACTCGAGAGTGCAGAAAAAATAGTACATGATTACAGCAAGAAATTGTTCTCTCGT 246
 QY 66 IleThrGluAspGlnLeuTrpArgAlaIleValTyrValTyrAspSerAlaPheHisProAsp 85
 DB 247 CTTACAGAAATGAATTTGTGAGAGCAAGATACATCTATGATTCAGCTTTTCATCTGAC 306
 QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
 DB 307 ACTGGTGAGAGATGATTTGTGATAGGAGAAATGTCAGCCAGGTTCCCATGAACATGACC 366
 QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
 DB 367 ATCACAGGTTGTATGATGACGTTTACAGGACTACGCCGGCTGTGCTGTTCTGGCAGTGG 426
 QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
 DB 427 ATTAACCAAGTCTTCATGCGCTGCTCAATTAACCAACAGAGAGTGGAGACGCCCTC 486
 QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165

DB 487 ACTGTCAATGAGTTGGGAACAGCTTACGTTCTGCAACAACTGGTGGCCGTAGCAACAGCT 546
 QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 DB 547 CTAGAGACTCAATGCATTTGACCAAGCATGTCTCACCACTGATAGACGTTTGTCTCCCTT 606
 QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
 DB 607 GCTGCCGTAGCTGCTGCTAATTCATTAATTAATTCATTAATGAGCAAGGAACCTCAAA 666
 QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
 DB 667 GTTGGCATTCCTCCGTCACGGATGAGAAATGGGAACCGCTTGGGGAGTCGCCGAACGCTGCG 726
 QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 DB 727 AAACAAGCCATCACGCAAGTGTGTCGTCCAGGATTCATGAGCCCTGGCATGGCC 786
 QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
 DB 787 ATCCCTCCATTCATTAATGAACACTTTGGAAAAGAACCTTTTGAAG 834
 RESULT 11
 ABN84365
 ID ABN84365 standard; cDNA; 1134 BP.
 XX
 AC ABN84365;
 XX
 DT 01-OCT-2002 (first entry)
 XX
 DE Alzheimer's disease associated protein ADPI-41 coding sequence.
 XX
 KW ADPI-4; human; neuropsychiatric disorder; neurological disorder;
 KW Alzheimer's disease; nootropic; neuroprotective; neuroleptic;
 KW antiparkinsonian; antidepressant; diagnosis; vaccine; human;
 KW sideroflexin 1; Sxfln1; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 11..979
 FT /*tag= a
 FT /product= "ADPI-41"
 FT
 XX MO200246221-A2.
 XX 13-JUN-2002.
 XX 10-DEC-2001; 2001WO-GB05459.
 XX 08-DEC-2000; 2000US-254431P.
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
 XX Herath HMA, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
 DR MPI: 2002-557575/59.
 DR P-PSDB; ABB79572.
 PT Isolated or recombinant polypeptide useful in the treatment of
 PT neuropsychiatric or neurological disorder comprises a specific amino
 PT acid sequence -
 XX
 XX Claim 6; Fig 2a; 51pp; English.
 XX
 CC The present sequence is the coding sequence of a human protein,
 CC named ADPI-41, which is differentially expressed (present in
 CC reduced amounts) in all areas of the brain tested except the
 CC amygdala in Alzheimer's disease (AD). The full-length open
 CC reading frame was amplified by PCR from human brain and liver
 CC cDNA libraries. ADPI-41 is useful as a marker and/or therapeutic
 CC target in AD. Its amino acid sequence is 95% identical to that of
 CC murine sideroflexin 1 (Sxfln1), recently identified in the mouse

CC model of sideroblastic anaemia. This is the first demonstration of
CC the human Sx1n protein and the first indication that it may be
CC associated with the occurrence of AD. ADPI-41 undergoes
CC alternative splicing to result in an isoform of 261 amino acids
CC (see ABB79573). The invention provides ADPI-41 polypeptides and
CC polynucleotides, vectors, host cells and antibodies. A claimed
CC method of screening for and/or diagnosis of a neuropsychiatric or
CC neurological disorder in a subject and/or monitoring the
CC effectiveness of therapy comprises detecting and/or quantifying
CC the amount of ADPI-41 polypeptide and/or nucleic acid molecule in a
CC sample from the subject. A claimed method for the prophylaxis
CC and/or treatment of a neuropsychiatric disorder or neurological
CC disorder comprises administering an ADPI-41 polypeptide or nucleic
CC acid. A method of screening for agents that modulate ADPI-41
CC activity or expression is also claimed. Disorders that may be
CC treated or diagnosed include AD, vascular dementia, Lewy body
CC dementia, schizophrenia, Parkinson's disease, multiple sclerosis,
CC and depression.

XX Sequence 1134 BP; 303 A; 268 C; 273 G; 287 T; 3 other;

Alignment Scores:
Pred. No.: 8.85e-114 Length: 1134
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 24 Gaps: 0

US-09-990-415A-2 (1-261) x ABN84365 (1-1134)

QY 6 GlyGlueuProleuAspIleasnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
DB 17 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAAAGCACTTCATT 76
QY 26 GlyArgAlaArgHisPheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
DB 77 GGACGAGCCCAATCATTTCTTCACTGTAACGTACCCAGAACCAATCTGTTAACCAACGAA 136
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValThrProGly 65
DB 137 CAAGTCGAGAGTGCAGAAAAAATAGTACATGATTACAGGCAAGAAATGTTCTCCTCGT 196
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProasp 85
DB 197 CTTACAGAAAAATGAATGTGGAGCAAGATACATCTATGATTCAGCTTTCATCCTGAC 256
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometasmetThr 105
DB 257 ACTGGTGAGAAGATGATTTGTATAGGAAGATGTACAGCCCGAGTCCCATGAACATGACC 316
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
DB 317 ATCAGACGGTGTATGATGACGCTTTTACAGGACTACGCCGGCTGTCTGTGGCAGTGG 376
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
DB 377 ATTAACCAAGTCTCTCAATGCCGTGCTCAATTACACCAACAGAGTGAGACGACCCCTTC 436
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
DB 437 ACTGTCAATGAGTGTGGAAACAGTTCAGTTTCTGCAACAACTGGTCCGTAGCAACAGCT 496
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
DB 497 CTAGGACTCAATGATGACCAAGCATGTCTCACCACAGTATAGGACGTTTGTCCCTTT 556
QY 186 AlaAlaValAlaAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
DB 557 GCTGCCGTAGCTGCTGTAATTCATTAATATTCATTAATGAGGCAAGGAACTCAA 616
QY 206 ValGlyIleProValAlaAspGluAlaGlyLysArgLeuGlyTyrSerValThrAlaAla 225

DB 617 GTTGCATTCCTCCGTACCGATGAGATGGAAACCGCTTGGGGAGTCGCCGAACGCTGCC 676
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
DB 677 AAACAAGCCATCAACGCAAGTGTGTCGTCAGGATTCATGCGACGCCCTGGCATGCCC 736
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
DB 737 ATCCCTCATTCATTAATGAACACTTGGAAAGAAGACCTTTTGAAG 784

RESULT 12

AAL26555
ID AAL26555 standard; cDNA; 1485 BP.

AC AAL26555;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 19012.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

PN WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.

PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.

PR 09-JUN-2000; 2000US-0211315.

PR 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer

PT Claim 1; Page 3556; 3695pp; English.

PS The invention relates to human breast cancer expressed polynucleotides

CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the

CC expression of certain markers and the cancerous state of breast cells.

CC The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterising treating and

CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic

XX activity.

SO Sequence 1485 BP; 375 A; 378 C; 365 G; 359 T; 8 other;

Alignment Scores:
Pred. No.: 1.31e-113 Length: 1485
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 22 Gaps: 0

US-09-990-415A-2 (1-261) x AAL26555 (1-1485)

QY 6 GlyGlueuProleuAspIleasnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
DB 112 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAAAGCACTTCATT 171

QY 26 G1YArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuLeuSerGlyAla 45
Db 172 GGACGAGCCAAATCATTTCTTCACTGTAACTGACCCAGAACATCTCTGTAAACCAAGAA 231
QY 46 GlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 232 CAACTCGAGAGTGCAGAAAAAATAGTACATGATTCACGCAAGGAATTGTTCCCTCTGGT 291
QY 66 IleThrGlnAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 292 CTTACAGAAATGAATGTGTGAGAGCAAAAGTACATCTATGATTCAGCTTTTCATCCTGAC 351
QY 86 ThrGlyGlnLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 352 ACTGGTGAGAGATGATTTTGATAGGAAGATGTCAGCCCAAGTCCCATGAACATGACC 411
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 412 ATCACAGGTTGTATGATGACGTTTACAGAGACTACGCGCGCTGTGCTGTCTGCGCAGTGG 471
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 472 ATTAACACAGTCTCTCAATGCCGCTGTCATATACACCAACAGAAGTGGAGACGCCCTC 531
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 532 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGCAACAACCTGTCGCCGTAGCAACAGCT 591
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 592 CTAGGACTCAATGCATGTGACCAAGCATGCTCACCACTGATAGGACGTTTGTCCCTTT 651
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
Db 652 GCTGCCGTAGCTGCTGCTAATTGCATTATATATTCATTAAATGAGGCAAGGAACTCAAA 711
QY 206 ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 712 GTTGCAATTCCTCCGTCACGGATGAGAATGGAAACCGCTGGGGAGTCGCGCAACGTCGG 771
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetalAlaIleProAlaMetala 245
Db 772 AAACAAGCCATCACGCAAGTGTCTGTCAGCATTCATGCGAAGCCCTGGCATGGCC 831
QY 246 IleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
Db 832 ATCCCTCCATTCATTTATGAACACTTTGGAAAAAGAAAGCCTTTTGAAG 879
RESULT 13
AAI61354
ID AAI61354 standard; cDNA; 2053 BP.
XX
AC AAI61354;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5343.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemoknetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.

XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB; AAM42198.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 5343; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemoknetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2053 BP; 582 A; 436 C; 451 G; 584 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,1e-113 Length: 2053
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 22 Gaps: 0
US-09-990-415A-2 (1-261) x AAI61354 (1-2053)
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 94 GGAGAACTACCAACCAACATTAACATCAAGAACCTCGATGGATCAAGCACTTTCATT 153
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuLeuSerGlyAla 45
Db 154 GGACGAGCCAAATCATTTCTTCACTGTAACTGACCCAGAACATCTCTGTAAACCAAGAA 213
QY 46 GlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 214 CAACTCGAGAGTGCAGAAAAAATAGTACATGATTCACGCAAGGAATTGTTCTCTGGT 273
QY 66 IleThrGlnAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 274 CTTACAGAAATGAATTTGTGAGAGCAAAAGTACATCTATGATTCAGCTTTTCATCCTGAC 333
QY 86 ThrGlyGlnLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 334 ACTGGTGAGAGATGATTTTGATAGGAAGATGTCAGCCCAAGTCCCATGAACATGACC 393

QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 394 ATCACAGGTTGATGATGACGTTTACAGGACTACGCCGGCTGTGCTGTCTGGCAGTGG 453
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSergLysAspThrProIle 145
Db 454 ATTAACCAAGTCTTCAATGCCGTCGTCATTAACCAACAGAGAGTGAGACGCCACCCCTC 513
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 514 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGCAACAACACTGTGCGGTAGCAACAGCT 573
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 574 CTAGAGCTCAATGCATGTGACCAAGCATGTCTCACCACTGATAGACGTTTGTCCCTTT 633
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db 634 GCTGCCGTAGCTGCTGCTAATTCATTAATATTCATTAATGAGGCAAGGAACTCAAA 693
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 694 GTTGGCATTCGCCGTACCGGATGAGATGGGAACCCGCTGGGGGAGTGGCGGAACGCTGCG 753
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 754 AAACAAGCCATCACGCAAGTGTGCTGCTCCAGGATTCATGAGCAGCCCTGGCATGGCC 813
QY 246 IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 814 ATCCCTTCATTCATTATGAAACACTTGGAAAAGAACCTTTTGAAG 861

RESULT 14

ABN83757
ID ABN83757 standard; cDNA; 2269 BP.
XX
AC ABN83757;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human protein cluster I gene.
XX
KW Protein cluster I; human; metabolic disorder; obesity; diabetes;
KW antidiabetic; diagnosis; therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 125..1093
FT /*tag= a
FT /product= "Protein cluster I"
XX
PN WO200242324-A1.
XX
PD 30-MAY-2002.
XX
PF 22-NOV-2001; 2001WO-SE02581.
XX
PR 24-NOV-2000; 2000SE-0004325.
XX
PA (PHAA) PHARMACIA AB.
XX
PI Attersand A;
XX
DR WPI; 2002-500277/53.
DR P-PSDB; ABB76448.
XX
PT Novel nucleic acid molecule encoding Protein Cluster I, useful in the
PT diagnosis of metabolic diseases, such as obesity and diabetes, and in
PT the identification of agents useful in the treatment of the diseases
XX
PS Claim 1; Page 30-32; 34pp; English.

CC The present invention relates to the identification of a human gene
CC family (see ABN83754-57) encoding a group of polypeptides (see
CC ABB76445-48) referred to as Protein Cluster I. This family of
CC homologous proteins was identified by an 'all-versus-all' BLAST
CC procedure using all Caenorhabditis elegans proteins in the
CC Wormpep20 database release. Proteins of unknown function were
CC compared to the Drosophila melanogaster flybase database, and
CC non-annotated protein clusters, conserved in both C. elegans and D.
CC melanogaster, were used in a BLAST procedure against the Celera
CC Human Genome Database, and Protein Cluster I proteins of unknown
CC function were selected for study. The human part of protein
CC cluster I comprises polypeptides encoded by 3 genes. The present
CC gene is expressed primarily in the immune system. The claimed
CC Protein Cluster I nucleic acid molecules and proteins are proposed
CC to be useful for differential identification of the tissue(s) or
CC cell type(s) present in a biological sample, for diagnosis of
CC diseases and disorders, including metabolic disorders and immune
CC disorders, especially obesity and diabetes, and for identifying
CC agents useful in the treatment of such diseases. The nucleic acid
CC molecules are also useful as hybridisation probes, for chromosome
CC and gene mapping, in PCR technologies, and in the production of
CC sense or antisense nucleic acids.

SQ Sequence 2269 BP; 631 A; 479 C; 506 G; 652 T; 1 other;

Alignment Scores:

Pred. No.: 2,42e-113 Length: 2269
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 24 Gaps: 0

US-09-990-415A-2 (1-261) x ABN83757 (1-2269)

QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db 131 GGAGAACTTACCACCAACATTAACTACAGCAAGCACTGATGGATCAAGCACTTTCATT 190
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSergLysAla 45
Db 191 GGACGAGCCCAATCATTTCTTCACTGTAACTGACCCAGGAACATTCGTTAAACCAAGAA 250
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 251 CAACTCGAGAGTGCAGAAATAATAGTACATGATTACAGCAGGAAATGTCTCCTCGGT 310
QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 311 CTTACAGAAATGAATGTGTGGAGAGCAAGATACATCTATGATTACGCTTTTCATCCTGAC 370
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 371 ACTGGTGAGAGATGATTTTGTAGGAGAAATGTCAAGCCAGGTTCCCATGAACATGACC 430
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db 431 ATCACAGGTTGATGATGACGTTTACAGGACTACGCCGGCTGTGCTGTGGCAGTGG 490
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSergLysAspThrProIle 145
Db 491 ATTAACCAAGTCTTCAATGCCGTCGTCATTAACCAACAGAGAGTGAGACGCCACCCCTC 550
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 551 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGCAACAACACTGTGCGGTAGCAACAGCT 610
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 611 CTAGAGCTCAATGCATGTGACCAAGCATGTCTCACCACTGATAGACGTTTGTCCCTTT 670
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205

Db 671 GCTGCCGTAGCTGCTGCTAATTCATTAATATTCATTAATGAGGCAAGGCACTCAAA 730
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 731 GTTGGCATTCCTCCGTACCGATGAGAAATGGGAACCGCTGGGGAGTCCGCGAAGCTGCG 790
QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db 791 AAACAAGCCATCACCAAGTTGCTGCTCCAGATTCATGCGACGCCCTGGCATGGCC 850
QY 246 IleProPheLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db 851 ATCCCTCATTCATTTATGAACACTTTGGAAAGAACCTTTTGAAG 898
RESULT 15
AAD21995
ID AAD21995 standard; cDNA; 2367 BP.
XX AAD21995;
AC AAD21995;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human transporters and ion channels (TRICH)-3 cDNA.
XX
KW Human; transporter and ion channel; TRICH; akinesia; cystic fibrosis;
KW diabetes mellitus; Parkinson's disease; myasthenia gravis; dementia;
KW cardiac disorder; angina; hypertension; myocarditis; hyperglycaemia;
KW neurological disorder; Alzheimer's disease; cataract; infertility;
KW Wilson's disease; schizophrenia; Grave's disease; addison's disease;
KW Huntington's disease; multiple sclerosis; meningitis; hypotensive;
KW cardiact; nootropic; neuroprotective; neuroleptic; ophthalmological;
KW antithyroid; anticonvulsant; goitre; antiinflammatory; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 125..1093
FT /*tag= a
FT /product= "Human transporters and ion channels
FT (TRICH)-3"
XX
PN WO200177174-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US11206.
XX
PR 06-APR-2000; 2000US-195595P.
PR 12-APR-2000; 2000US-196872P.
PR 20-APR-2000; 2000US-199020P.
PR 28-APR-2000; 2000US-200552P.
PR 05-MAY-2000; 2000US-202348P.
PR 11-MAY-2000; 2000US-203495P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Reddy R, Thornton M, Borowsky ML, Tang YT, Khan FA, Tribouley CM;
PI Gandhi AR, Yao MG, Sanjanwala MS, Baughn MR, Nguyen DB;
PI Policky JL, Yue H, Selhamer JJ, Walla NK, Lal P, Kearney L;
PI Walsh RT, Lu DAM, Lu Y, Greene BD, Raumann BE, Patterson C;
XX
DR WPI; 2002-017448/02.
DR P-PSDB; AAE13276.
XX
XX Polypeptides of human transporters and ion channels, useful for
PT diagnosing, treating or preventing disorders of transport,
PT neurological, muscle, immunological and cell proliferative disorders -
XX
XX Claim 5; Page 139-140; 150pp; English.
XX
XX The invention relates to human transporters and ion channels (TRICH)
CC and the polynucleotides encoding them. The composition comprising TRICH
CC or agonist of TRICH is useful for treating a disease or condition

CC associated with decreased expression of functional TRICH or condition
CC associated with overexpression of TRICH respectively. The composition
CC comprising Ab is useful for diagnosing a condition of disease associated
CC with expression of TRICH in a subject, where the disorders include a
CC transport disorder such as akinesia, cystic fibrosis, diabetes mellitus,
CC Parkinson's disease, myasthenia gravis, cardiac disorders associated
CC with transport e.g. angina, hypertension, myocarditis, neurological
CC disorders associated with transport e.g. Alzheimer's disease, Wilson's
CC disease, schizophrenia, cataracts, infertility, hyperglycaemia, Grave's
CC disease, goitre, addison's disease, Huntington's disease, dementia,
CC multiple sclerosis, bacterial and viral meningitis. TRICH DNA is useful
CC for generating a transcript image of a tissue or cell type, which
CC represents the global pattern of gene expression by a particular tissue
CC or cell type and for analysing the proteome of a tissue or cell type.
CC TRICH DNA is used in gene therapy. The present sequence is human
CC TRICH3 cDNA.
XX
SQ Sequence 2367 BP; 644 A; 502 C; 530 G; 691 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,58e-113 Length: 2367
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
DB: 24 Gaps: 0
US-09-990-415A-2 (1-261) x AAD21995 (1-2367)
QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrAspGlnSerThrPheLeu 25
Db 131 GGAGAACCTACCAACCAACATTAAACATCAAGAACCTCGATGGGATCAAGCACTTCATT 190
QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
Db 191 GGACGAGCCATCATTTCTTCACTGTAAGTCAAGCAAGCAAGCAATCTGTTAAACCAAGAA 250
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db 251 CAACTCGAGAGTGGCAGAAATAAGTACATGATTACAGGAGGAGGAATGTTCTCCTCGGT 310
QY 66 IleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db 311 CTTACAGAAATGAATGTTGTGAGAGCAAGTACATCATGATTCAGTTTCATCCTGAC 370
QY 86 ThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
Db 371 ACTGGTGAAGATGATTTGTATGAGAGAATGTCAAGCCAGGTTCCCATGAACATGACC 430
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTyr 125
Db 431 ATCACAGGTGTATGATGACGTTTACAGACTACGCCGCGTGTCTGTGCGAGTGG 490
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db 491 ATTAACCAAGTCTTCAATGCCGTCGTCATTAACCAACAGAGTGAAGACGACCCCTC 550
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db 551 ACTGTCAATGAGTTGGGAACAGCTTACGTTTCTGCAACAACACTGTCCTGACACAGCT 610
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db 611 CTAGCACTCAATGCAATGACCAACATGCTCTCACCACTGATGAGACGTTTGTCCCTTT 670
QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db 671 GCTGCCGTAGCTGCTGCTAATTCATTAATATTCATTAATGAGGCAAGGCAACTCAAA 730
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db 731 GTTGGCATTCCTCCGTACCGATGAGAAATGGGAACCGCTGGGGAGTCCGCGAAGCTGCG 790

QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
||||| ||| |||||||:::||||| ||| ||| |||
Db 791 AAACAAGCCATCAGCAAGTGTGCTGCCAGATTCATGGCAGCCCTGGCATGGCC 850
QY 246 IleProProlleuIleMetAspThrleuGluLysLysAspPheleuLys 261
||||| ||| |||||||:::||||| ||| ||| |||
Db 851 ATCCCTCCATTATGAACTTTGGAAAAGAAAGCCCTTTTGAAG 898

Search completed: October 16, 2003, 16:10:52
Job time : 270 secs

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-990-415A-1 (1-1232)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
DB 450 ATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAACCTCGCTGGAC 509
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 510 CAAAGTACTTCTCTGGCAGAGCCCGGCACTTTTCACTGTACTGATCCTCGAAATCTG 569
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 570 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGC 629
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
DB 630 GTGGTGAACCCAGGGATCACCGAGACCACTGTGAGGGCCCAAGTATGTATGACTCC 689
QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB 690 GCCTTCCATCCGGACACAGGGGAGAGGTGCTCTGATGGCCGATGTCAAGCCAGGTG 749
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 750 CCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGAGAACCCCAACCGTG 809
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 810 GTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCCATGTTAATTACTCCAAACCGCAGT 869
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 870 GGTGACACTCCCATCACTGTGAGGAGCTGGGACAGCCTATGTGAGTGCCACCACTGGA 929
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 930 GCTGTGGCCACGGCCCTGGGACCTCAATCCCTCACCAAGCACCTGCCCTTGTCGGC 989
QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 990 AGATTGTGCCCTTTGCAGCAGTGGCAGCTGCCAATGCATCAACATCCCCCTGATGAGG 1049
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 1050 CAGAGAGAGCTGCAGGTGGGCAATCCGGTGCTGATGAGGCAAGCTTGCGCTAG 1109
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 1110 TCGGTGACTGCAGCCCAAGCAGGAATCTTCCAGGTGTGATTTCAAGAAATCTGCATGGCG 1169
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
DB 1170 ATTCTGCCATGGCCATCCCACTGATGACACTCTGAGAGAAAGAACTTCTCTG 1229
QY 261 Lys 261
DB 1230 AAG 1232

RESULT 2
US-09-867-550-1343

; Sequence 1343, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867, 550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1343
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2)
; OTHER INFORMATION: wherein any n is one of a or t or c or g
US-09-867-550-1343

Alignment Scores:
Pred. No.: 3,79e-171 Length: 1375
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-867-550-1343 (1-1375)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
DB 91 ATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGAACCTCGCTGGAC 150
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 151 CAAAGTACTTCTCTGGCAGAGCCCGGCACTTTTCACTGTACTGATCCTCGAAATCTG 210
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 211 CTGCTGTCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTGCAAGACTACAGGGCCGC 270
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
DB 271 GTGGTGAACCCAGGGATCACCGAGACCACTGTGAGGGCCCAAGTATGTATGACTCC 330
QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB 331 GCCTTCCATCCGGACACAGGGGAGAGGTGCTCTGATTTGCCCATGTCAAGCCAGGTG 390
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 391 CCCATGAACATGACCATCACTGGCTGCATGCTCACATTTCTACAGAGAACCCCAACCGTG 450
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 451 GTGTTTGGCAGTGGGTGAATGCTCTTCAATGCCATGTTAATCACTCAACCGCAGT 510
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 511 GGTGACACTCCCATCACTGTGAGGCAAGCTGGGACAGCCATGTGAGTCCACCACTGGA 570
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 571 GCTGTGGCCACGGCCCTGGGACTCAATCCCTCACCAAGCAAGCTGCCCTTGTCGGC 630
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 631 AGATTGTGCCCTTTGCAGCAGTGGCAAGCTGCCAATGCATCAACATCCCTGATGAGG 690
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 691 CAGAGAGAGCTGCAGGTGGGCAATCCCGGTGCTGATGAGGAGGTCAAGGCTTGCTTAC 750
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240

```
Db 751 TCGGTGACTGCAGCCAGCAGGGAATCTCCAGGTGTGATTTCAAGAACTCGCATGGCG 810
Oy 241 ILeProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 811 ATTCCTGCCATGGCCATCCACCACATGATCATGACACTCTGGAGAGAAGAACTTCCTG 870
Oy 261 Lys 261
Db 871 AAG 873

RESULT 3
US-10-037-270-1016
; Sequence 1016, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aldong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL-genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (247)..(1371)
US-10-037-270-1016

Alignment Scores:
Pred. No.: 5.45e-171 Length: 1729
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-990-415A-2 (1-261) x US-10-037-270-1016 (1-1729)
Oy 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 247 ATGGAAGCAAAATGGGTGAATTGCTTTAGACATCAACATCCAGGAACCTCGCTGGGAC 306
Oy 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 307 CAAAGTACTTCTCTGGGAGAGCCCGGCACTTTTCACTGTACTGATCCTCGAAATCTG 366
Oy 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
```

```
Db 367 CTGCTGTCCGGGGCCACAGCTGGAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCCGGC 426
Oy 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValTyrAspSer 80
Db 427 GTGGTGACCCAGGGATCACCAGGACCAAGCTGTGGAGGGCCAAAGTATGTATGACTCC 486
Oy 81 AlaPheHisProAspThrGlyGluLysValIleuIleGlyArgMetSerAlaGlnVal 100
Db 487 GCCTTCCATCCGGACACACAGGGAGAAAGTGGTCTGATTTGGCCGCAATGCACCCAGTG 546
Oy 101 PrometaSmMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 547 CCCATGACATGACCATCAGTGGCTGCATGCTCACATTTACAGAGAACCCCAACCGTG 606
Oy 121 ValPheTyrPcIntTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 607 GTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCCATTTGTTAACTACTCCAACCGCAGT 666
Oy 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrgly 160
Db 667 GGTGACACTCCCATCACTGTGAGGAGCAGCTGGGGACAGCCTATGTAGTGCACCACTGGA 726
Oy 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 727 GCTGTGGCCAGGCGCCCTGGGACTCAAAATCCCTCACCAAGCACCTGCCCTTGGTGGCG 786
Oy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 787 AGATTGTGCGCTTTTGACAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCGCTGATGAGG 846
Oy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 847 CAGAGAGAGCTGCAGGTGGGCATCCCGGTGGCTGATGAGGCAGGTCAAGGCTTGACTAC 906
Oy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 907 TCGGTACTGCAGCCCAAGCAGGGAATCTTCCAGGTGATTTCAAGAAATCTGCATGGCG 966
Oy 241 ILeProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 967 ATTCCTGCCATGGCCATCCACCACACTGATCATGACACTCTGGAGAGAAGAACTTCCTG 1026
Oy 261 Lys 261
Db 1027 AAG 1029

RESULT 4
US-10-014-338-1
; Sequence 1, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADPI-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENAT
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1121)..(1122)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1125)..(1126)
; OTHER INFORMATION: where "n" is any nucleotide
```

RESULT 4

US-9-867-550-1343

Sequence 1343, Application US/09867550

Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.

APPLICANT: Mehrahan, Fuad,

APPLICANT: Conley, Pamela

APPLICANT: Law, Debbie

APPLICANT: Topper, James

TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

TITLE OF INVENTION: Thereby

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867,550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for windows Version 4.0

SEQ ID NO 1343

LENGTH: 1375

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(2)

OTHER INFORMATION: wherein any n is one of a or t or c or g

US-09-867-550-1343

Query Match 70.6%; Score 869.4; DB 9; Length 1375;
Best Local Similarity 99.9%; Pred. No. 4.6e-253;
Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	362	ACGCGTGACGTCCCGCGTGATGGCTGGGAGGCGCGGCGAGACAGGAGGAGAGG	421
DB	3	ACGCGTGACGTCCCGCGTGATGGCTGGGAGGCGCGGCGAGACAGGAGGAGAGG	62
OY	422	AAGCGGTTCTGAGAGCTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	481
DB	63	AAGCGGTTCTGAGAGCTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	122
OY	482	CATCAACATCCAGGAACTCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	541
DB	123	CATCAACATCCAGGAACTCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	182
OY	542	TTTCACTGTACTGATCCTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	601
DB	183	TTTCACTGTACTGATCCTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	242
OY	602	GAACATCGTGCAGAACTACAGGCGCGCGCTGGTGACCCCAAGGATCAACGAGCAGCT	661
DB	243	GAACATCGTGCAGAACTACAGGCGCGCGCTGGTGACCCCAAGGATCAACGAGCAGCT	302
OY	662	GTGAGGCGCAAGTATGTATGATGATGATGATGATGATGATGATGATGATGATG	721
DB	303	GTGAGGCGCAAGTATGTATGATGATGATGATGATGATGATGATGATGATGATG	362
OY	722	CCTGATGGCGCGCATGTGACGCCAGGCTCCATGAAATGACATGACATGACATG	781
DB	363	CCTGATGGCGCGCATGTGACGCCAGGCTCCATGAAATGACATGACATGACATG	422
OY	782	CACATCTACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	841
DB	423	CACATCTACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	482
OY	842	TGCCATGTGTAATCTACCAACCGAGTGGTGACACTCCCATCACTGTGAGGAGCTGG	901
DB	483	TGCCATGTGTAATCTACCAACCGAGTGGTGACACTCCCATCACTGTGAGGAGCTGG	542
OY	902	GACAGCCTATGTAGTGCCACACATGGAGCTGTGGCCACGCGCTGGGAGCTCAATCCCT	961
DB	543	GACAGCCTATGTAGTGCCACACATGGAGCTGTGGCCACGCGCTGGGAGCTCAATCCCT	602

OY	962	CACCAAGCACCCTGCCCCCTTGCTGGCAGATTTGTGCCCTTTCAGACAGTGAGCTGC	1021
DB	603	CACCAAGCACCCTGCCCCCTTGCTGGCAGATTTGTGCCCTTTCAGACAGTGAGCTGC	662
OY	1022	CAACTGCATCAACATCCCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1081
DB	663	CAACTGCATCAACATCCCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	722
OY	1082	TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1141
DB	723	TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	782
OY	1142	GGTGTGATTTCAAGAAATCTGCATGCGGATTCCTGCCATGGCCATCCACACTGAT	1201
DB	783	GGTGTGATTTCAAGAAATCTGCATGCGGATTCCTGCCATGGCCATCCACACTGAT	842
OY	1202	GGACACTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1232
DB	843	GGACACTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	873

RESULT 5

US-09-990-415A-7

Sequence 7, Application US/09990415A

Patent No. US20020165182A1

GENERAL INFORMATION:

APPLICANT: Pharmacia AB

TITLE OF INVENTION: Protein Cluster I

FILE REFERENCE: 00349

CURRENT APPLICATION NUMBER: US/09/990,415A

CURRENT FILING DATE: 2001-11-21

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patent version 3.0

SEQ ID NO 7

LENGTH: 2269

TYPE: DNA

ORGANISM: human

FEATURE:

NAME/KEY: CDS

LOCATION: (125)..(1093)

NAME/KEY: misc_feature

LOCATION: (25)..(25)

OTHER INFORMATION: n-A,T,G or C

US-09-990-415A-7

Query Match 32.4%; Score 399.4; DB 10; Length 2269;
Best Local Similarity 66.6%; Pred. No. 1.5e-110;
Matches 571; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

OY	376	GCGTGTGCTGGGAGGCGCGCGCGCGAGCAGGAGGAGGAGGAGGAGGAGGAGG	435
DB	42	GCGTGTGCTGGGAGGCGCGCGCGCGAGCAGTGAAGCGGAGGAGGAGGAGGAGG	101
OY	436	AGCTTCAGAGAGGATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCAGG	495
DB	102	GCGGCGGGAAGCGTCCGGACCATGTCTGAGAACTACCAACCAATTAACATCAGG	161
OY	496	AACCTGCTGGGAGCAAAAGTACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGG	555
DB	162	AACCTGATGGGATCAAAAGCATTTTATTTGAGAGGAGGAGGAGGAGGAGGAGG	221
OY	556	ATCTCGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	615
DB	222	ATCTCGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	281
OY	616	ACTACAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	675
DB	282	ACTACAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	341
OY	676	ATGTGTATGACTCCGCTTCATCCGAGACACAGGAGGAGGAGGAGGAGGAGGAGG	735
DB	342	ATGTGTATGACTCCGCTTCATCCGAGACACAGGAGGAGGAGGAGGAGGAGGAGG	401